(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 December 2020 (10.12.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/245376 A1

(51) International Patent Classification:
*C07K 14/435* (2006.01) *C07K 14/43* (2006.01)

(21) International Application Number:
PCT/EP2020/065652

(22) International Filing Date:
05 June 2020 (05.06.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
19178676.3 06 June 2019 (06.06.2019) EP

(71) Applicants: **NOMAD BIOSCIENCE GMBH** [DE/DE]; Türkenstr. 16, 80333 Munich (DE). **UAB NOMADS** [LT/LT]; Geležinio Vilko street 29A, 01112 Vilnius (LT).

(72) Inventors: **DENKOVSKIENE, Erna**; S. Staneviciaus g. 55-23, 07104 Vilnius (LT). **MISIUNAS, Audrius**; Sviliškiu g. 4-16, 06152 Vilnius (LT). **RAŽANSKIENE, Aušra**; Arkos g. 7-7, Klevines vs, 14180 Vilniaus raj. (LT).

(74) Agent: **WÄCHTERSHÄUSER & HARTZ PATENTANWALTSPARTNERSCHAFT MBB**; Weinstr. 8, 80333 Munich (DE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Published:**
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

WO 2020/245376 A1

(54) Title: KLEBICINS FOR THE CONTROL OF KLEBSIELLA

(57) Abstract: The invention provides a protein having cytotoxic activity against *Klebsiella*, said protein having a lipid ll-cleaving activity or a pore-forming capability in a cell membrane of *Klebsiella*.

# Klebicins for the control of *Klebsiella*

## FIELD OF THE INVENTION

The present invention provides proteins having cytotoxic activity against *Klebsiella*. The invention also provides compositions, including pharmaceutical compositions, comprising one or more of said proteins. The invention further provides a protein having cytotoxic activity against *Klebsiella*, and a composition comprising the protein, for use in therapy. Also provided is a protein, or composition comprising the protein, for use in a method of treating infection of a subject with *Klebsiella*. Further provided is an oral enteric formulation for delivering the protein or composition to the small or large intestine. Further provided is an pulmonary formulation for delivering the protein or composition to the lungs. Also provided is a method of preventing or reducing infection or contamination of an object with Klebsiella, a method of treating infection with Klebsiella of a subject or patient in need thereof, and a process of producing a composition comprising the protein. The invention also provides a process of producing a composition the protein of the invention. The invention further provides a nucleic acid molecule encoding a protein having cytotoxic activity against *Klebsiella*; a plant, plant tissue or plant cell comprising the protein; and a plant, plant tissue or plant comprising the nucleic acid molecule.

## BACKGROUND OF THE INVENTION

Klebsiellae are nonmotile, rod-shaped, gram-negative bacteria, encased by the polysaccharide capsule providing resistance against many host defense mechanisms. Klebsiellae are opportunistic pathogens found in the environment and in mammalian mucosal surfaces. Three species in the genus *Klebsiella* are commonly associated with illness in humans: *K. pneumoniae, K. oxytoca,* and *K. granulomatis*. Recently it was discovered that two more *Klebsiella* species, *K. variicola* and *K. quasipneumoniae* also can cause deadly infections (Long et al. 2017).

The principal pathogenic reservoirs of infection are the gastrointestinal tract of patients and the hands of hospital personnel. Outside hospitals, infection with *Klebsiella* typically occurs in the lungs. The illness typically affects middle-aged and older men with debilitating diseases such as alcoholism, diabetes, or chronic bronchopulmonary disease (Chan et al., 2009). This patient population is believed to have impaired respiratory host defenses. The organisms gain access after the host aspirates colonizing oropharyngeal microbes into the lower respiratory tract (Hirsche et al., 2005).

In recent years, Klebsiellae have become important pathogens in nosocomial infections. Common sites of nosocomial infections include the urinary tract, lower respiratory tract, biliary tract, and surgical wound sites. The spectrum of clinical syndromes includes pneumonia,

bacteremia, thrombophlebitis, urinary tract infection (UTI), cholecystitis, diarrhea, upper respiratory tract infection, wound infection, osteomyelitis, and meningitis (Miftode et al., 2008). The presence of invasive devices, contamination of respiratory support equipment, use of urinary catheters, and use of antibiotics are factors that increase the likelihood of nosocomial infection with *Klebsiella* species (Weisenberg et al., 2009).

*K. pneumoniae* is one of the six pathogens causing hospital ESKAPE (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, Enterobacter*) infections, which readily develop resistance to antibiotics. In 2016, an outbreak of nosocomial pneumonia was reported in China, where five surgical patients died from infection with hypervirulent carbapenem resistant *K. pneumoniae* (CRPK) ST11 strain (Gu et al, 2017). Such strain can be named as "superbug" as it is hypervirulent and hyper-resistant to antibiotics. ST11 CR-HvKP strains are infecting relatively healthy people with normal immunity. These are mucoid strains, sticking to all surfaces in Intensive Care Units. Colistin, which is a last resort antibiotic against carbapenem resistant enterobacteriae, has little effect against such strains. For the time being, ceftazidime/avibactam can be used to treat such infections, but the resistance to these antibiotics can be acquired also very soon.

Ever increasing pathogen drug resistance is a global problem and the development of a new generation of antimicrobial substances is urgently needed. Bacteria, competing with each other for ecological niches, produce toxic proteins, called bacteriocins. Bacteriocins usually kill only closely related bacteria, belonging to the same species or genera. Their mechanisms of action are diverse, including pore-formation, DNase and RNase activities, inhibition of protein synthesis or DNA replication, etc. The bacteriocins produced by Gram-positive bacteria are usually called bacteriocins of a determined class depending on their properties, while bacteriocins produced by Gram-negative strains are divided into the colicin-type bacteriocins (high molecular mass, 25-80,000 Da) or microcins (low molecular mass, <10,000 Da) (Lagos et al, 2009).

Antimicrobial peptides are not only produced by bacteria, but also all kinds of organisms when confronted with bacterial infection. Antimicrobial peptides are used in medicine as peptide antibiotics: colistin (polymyxin from *Paenibacillus polymyxa*), vancomycin (from *Amycolatopsis orientalis*). However, such antibiotics are mostly used for topical applications or as last resort drugs. Another problem with peptides is inefficient and costly purification from natural sources. It can be overcome by chemical synthesis, however, it is also expensive. Production of recombinant peptides in heterologous hosts is also difficult because of toxicity to host cells (Li, 2011).

There are no registered colicin-like antibiotics to date. However, there are studies in scientific literature where colicin-like bacteriocins are used as potential antimicrobials against Gram-negative pathogens. Most studied are colicins, several research groups are working with pyocins (Grinter et al., 2013; Ghequire, de Mot, 2014). For the time being, bacteriocins from *Klebsiella* have received very little attention and only few studies are published (James et al, 1987; Riley et al, 2001; Chavan et al, 2005) that belong to the nuclease class. Detailed studies of expression, purification and activity tests of klebicins are lacking.

Departing from the prior art, it is an object of the invention to provide an agent that is active against *Klebsiella*. Is also an object of the invention to provide an agent or composition that can be used to treat *Klebsiella* infections of a subject, notably infections by antibiotic-resistant *Klebsiella*. It is a further object to provide a method of preventing or reducing contamination of an object such as food with one or more *Klebsiella* species.

## SUMMARY OF THE INVENTION

The inventors have found novel bacteriocins that are active against *Klebsiella*. Thus, the invention provides the following.

1) A protein having cytotoxic activity against *Klebsiella*, said protein preferably having a lipid II-cleaving activity or a pore-forming capability in a cell membrane of *Klebsiella* cells.

2) The protein according to item 1, comprising or consisting of a first amino acid sequence segment and a second amino acid sequence segment, wherein the first amino acid sequence segment is capable of binding to components of *Klebsiella* cells and the second amino acid sequence segment has a lipid II-cleaving activity or a pore-forming capability in a cell membrane of *Klebsiella* cells.

3) A protein having cytotoxic activity against *Klebsiella*, said protein comprising or consisting of a first amino acid sequence segment and a second amino acid sequence segment, wherein the first segment is preferably the N-terminal segment of said protein and the second segment is the C-terminal segment of said protein.

4) The protein according to any one of items 2 or 3,

(A) wherein the first segment comprises or consists of the amino acid sequence of

- (A-i) from amino acid residue 1 to 128 of SEQ ID NO: 1 (KpneM),
- (A-ii) from amino acid residue 1 to 127 of SEQ ID NO: 2 (KvarM),
- (A-iii) from amino acid residue 1 to 123 of SEQ ID NO: 3 (KpneM2),
- (A-iv) from amino acid residue 1 to 118 of SEQ ID NO: 4 (KaerM),
- (A-v) from amino acid residue 1 to 170 of SEQ ID NO: 5 (KpneA),
- (A-vi) from amino acid residue 1 to 172 of SEQ ID NO: 6 (KaerA),

(A-vii) from amino acid residue 1 to 255 of SEQ ID NO: 7 (Koxy),

(A-viii) from amino acid residue 1 to 288 of SEQ ID NO: 8 (Kpnela), or

(A-ix) from amino acid residue 1 to 236 of SEQ ID NO: 9 (Kvarla);

or

(B) wherein the first segment comprises an amino acid sequence

(B-i) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 128 of SEQ ID NO: 1,

(B-ii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 127 of SEQ ID NO: 2,

(B-iii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 123 of SEQ ID NO: 3,

(B-iv) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 118 of SEQ ID NO: 4,

(B-v) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 170 of SEQ ID NO: 5,

(B-vi) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 172 of SEQ ID NO: 6,

(B-vii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 255 of SEQ ID NO: 7,

(B-viii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 288 of SEQ ID NO: 8, or

(B-ix) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 236 of SEQ ID NO: 9;

or

(C) wherein the first segment comprises an amino acid sequence

(C-i) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 128 SEQ ID NO: 1,

(C-ii) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 127 SEQ ID NO: 2,

(C-iii) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 123 of SEQ ID NO: 3,

(C-iv) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 118 of SEQ ID NO: 4,

(C-v) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 170 SEQ ID NO: 5,

(C-vi) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 172 of SEQ ID NO: 6,

(C-vii) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 255 SEQ ID NO: 7,

(C-viii) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 288 SEQ ID NO: 8, or

(C-ix) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 236 SEQ ID NO: 9.

5) The protein according to any one of items 2 or 3,

(A) wherein the amino acid sequence of the first segment is the amino acid sequence of

(A-i) from amino acid residue 1 to 128 of SEQ ID NO: 1 (KpneM),

(A-ii) from amino acid residue 1 to 127 of SEQ ID NO: 2 (KvarM),

(A-iii) from amino acid residue 1 to 123 of SEQ ID NO: 3 (KpneM2),

(A-iv) from amino acid residue 1 to 118 of SEQ ID NO: 4 (KaerM),

(A-v) from amino acid residue 1 to 170 of SEQ ID NO: 5 (KpneA),

(A-vi) from amino acid residue 1 to 172 of SEQ ID NO: 6 (KaerA),

(A-vii) from amino acid residue 1 to 255 of SEQ ID NO: 7 (Koxy),

(A-viii) from amino acid residue 1 to 288 of SEQ ID NO: 8 (Kpnela), or

(A-ix) from amino acid residue 1 to 236 of SEQ ID NO: 9 (Kvarla);

or

(B) wherein the amino acid sequence of the first segment has

(B-i) at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 128 of SEQ ID NO: 1,

(B-ii) at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 127 of SEQ ID NO: 2,

(B-iii) at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 123 of SEQ ID NO: 3,

(B-iv) at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 118 of SEQ ID NO: 4,

(B-v) at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 170 of SEQ ID NO: 5,

(B-vi) at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 172 of SEQ ID NO: 6,

(B-vii) at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 255 of SEQ ID NO: 7,

(B-viii) at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 288 of SEQ ID NO: 8, or

(B-ix) at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 236 of SEQ ID NO: 9;

or

(C) wherein the amino acid sequence of the first segment has

(C-i) from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 128 of SEQ ID NO: 1,

(C-ii) from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 127 of SEQ ID NO: 2,

(C-iii) from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 123 of SEQ ID NO: 3,

(C-iv) from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 118 of SEQ ID NO: 4,

(C-v) from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 170 of SEQ ID NO: 5,

(C-vi) from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 172 of SEQ ID NO: 6,

(C-vii) from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 255 of SEQ ID NO: 7,

(C-viii) from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 288 of SEQ ID NO: 8, or

(C-ix) from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 236 of SEQ ID NO: 9.

6) The protein according to item 4 or 5, wherein in item (B) any one of the sequence identities is at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, even more preferably at least 95%, and most preferably at least 97%; and/or

wherein in item (C) the number of said amino acid substitutions, additions, insertions and/or deletions is from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10, and most preferably at least 1 to 5 compared to any one of said amino acid sequences.

7) The protein according to any one of items 2 to 6,

(D) wherein the second segment comprises or consists of the amino acid sequence of

(D-i) from amino acid residue 129 to 278 of SEQ ID NO: 1 (KpneM),

(D-ii) from amino acid residue 128 to 276 of SEQ ID NO: 2 (KvarM),

(D-iii) from amino acid residue 124 to 272 of SEQ ID NO: 3 (KpneM2),

(D-iv) from amino acid residue 119 to 266 of SEQ ID NO: 4 (KaerM),

(D-v) from amino acid residue 171 to 377 of SEQ ID NO: 5 (KpneA),

(D-vi) from amino acid residue 173 to 379 of SEQ ID NO: 6 (KaerA),

(D-vii) from amino acid residue 256 to 452 of SEQ ID NO: 7 (Koxy),

(D-viii) from amino acid residue 289 to 466 of SEQ ID NO: 8 (Kpnela), or

(D-ix) from amino acid residue 237 to 414 of SEQ ID NO: 9 (Kvarla);

or

(E) wherein the second segment comprises an amino acid sequence

(E-i) having at least 70% sequence identity to the amino acid sequence from amino acid residue 129 to 278 of SEQ ID NO: 1,

(E-ii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 128 to 276 of SEQ ID NO: 2,

(E-iii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 124 to 272 of SEQ ID NO: 3,

(E-iv) having at least 70% sequence identity to the amino acid sequence from amino acid residue 119 to 266 of SEQ ID NO: 4,

(E-v) having at least 70% sequence identity to the amino acid sequence from amino acid residue 171 to 377 of SEQ ID NO: 5,

(E-vi) having at least 70% sequence identity to the amino acid sequence from amino acid residue 173 to 379 of SEQ ID NO: 6,

(E-vii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 256 to 452 of SEQ ID NO: 7,

(E-viii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 289 to 466 of SEQ ID NO: 8, or

(E-ix) having at least 70% sequence identity to the amino acid sequence from amino acid residue 237 to 414 of SEQ ID NO: 9;

or

(F) wherein the second segment comprises an amino acid sequence

(F-i) having from 1 to 30 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 129 to 278 of SEQ ID NO: 1,

(F-ii) having from 1 to 30 amino acid substitutions, additions, insertions or deletions compared to the amino acid sequence of from amino acid residue 128 to 276 of SEQ ID NO: 2,

(F-iii) having from 1 to 30 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 124 to 272 of SEQ ID NO: 3,

(F-iv) having from 1 to 30 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 119 to 266 of SEQ ID NO: 4,

(F-v) having from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 171 to 377 of SEQ ID NO: 5,

(F-vi) having from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 173 to 379 of SEQ ID NO: 6,

(F-vii) having from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 256 to 452 of SEQ ID NO: 7,

(F-viii) having from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 289 to 466 of SEQ ID NO: 8,

(F-ix) having from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 237 to 414 of SEQ ID NO: 9.

8) The protein according to any one of items 2 to 7,

(D) wherein the amino acid sequence of the second segment comprises or consists of the amino acid sequence of

(D-i) from amino acid residue 129 to 278 of SEQ ID NO: 1 (KpneM),

(D-ii) from amino acid residue 128 to 276 of SEQ ID NO: 2 (KvarM),

(D-iii) from amino acid residue 124 to 272 of SEQ ID NO: 3 (KpneM2),

(D-iv) from amino acid residue 119 to 266 of SEQ ID NO: 4 (KaerM),

(D-v) from amino acid residue 171 to 377 of SEQ ID NO: 5 (KpneA),

(D-vi) from amino acid residue 173 to 379 of SEQ ID NO: 6 (KaerA),

(D-vii) from amino acid residue 256 to 452 of SEQ ID NO: 7 (Koxy),

(D-viii) from amino acid residue 289 to 466 of SEQ ID NO: 8 (Kpnela), or

(D-ix) from amino acid residue 237 to 414 of SEQ ID NO: 9 (Kvarla)

or

(E) the amino acid sequence of the second segment has

(E-i) at least 80% sequence identity to the amino acid sequence from amino acid residue 129 to 278 of SEQ ID NO: 1,

(E-ii) at least 80% sequence identity to the amino acid sequence from amino acid residue 128 to 276 of SEQ ID NO: 2,

(E-iii) at least 80% sequence identity to the amino acid sequence from amino acid residue 124 to 272 of SEQ ID NO: 3,

(E-iv) at least 80% sequence identity to the amino acid sequence from amino acid residue 119 to 266 of SEQ ID NO: 4,

(E-v) at least 80% sequence identity to the amino acid sequence from amino acid residue 171 to 377 of SEQ ID NO: 5,

(E-vi) at least 80% sequence identity to the amino acid sequence from amino acid residue 173 to 379 of SEQ ID NO: 6,

(E-vii) at least 80% sequence identity to the amino acid sequence from amino acid residue 256 to 452 of SEQ ID NO: 7,

(E-viii) at least 80% sequence identity to the amino acid sequence from amino acid residue 289 to 466 of SEQ ID NO: 8, or

(E-ix) at least 80% sequence identity to the amino acid sequence from amino acid residue 237 to 414 of SEQ ID NO: 9,

or

(F) wherein the amino acid sequence of the second segment has

(F-i) from 1 to 30 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 129 to 278 SEQ ID NO: 1,

(F-ii) from 1 to 30 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 128 to 276 SEQ ID NO: 2,

(F-iii) from 1 to 30 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 124 to 272 of SEQ ID NO: 3,

(F-iv) from 1 to 30 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 119 to 266 of SEQ ID NO: 4,

(F-v) from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 171 to 377 SEQ ID NO: 5,

(F-vi) from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 173 to 379 of SEQ ID NO: 6,

(E-vii) from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 256 to 452 SEQ ID NO: 7,

(F-viii) from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 289 to 466 SEQ ID NO: 8, or

(F-ix) from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 237 to 414 SEQ ID NO: 9.

9) The protein according to any one of item 7 or 8, wherein said first segment is any one of items A-i to A-iv, B-i to B-iv, or C-i to C-iv and said second segment is any one of items D-i to D-iv, E-i to E-iv, or F-i to F-iv.

10) The protein according to item 9, wherein said first segment is any one or more of items A-i to A-iv and said second segment is any one of items D-i to D-iv, respectively; or said first segment is any one of items B-i to B-iv and said second segment is any one of items E-i to E-iv, respectively; or said first segment is any one of C-i to C-iv and said second segment is any one of items F-i to F-iv, respectively.

11) The protein according to any one of item 7 or 8, wherein said first segment is any one of items A-v to A-ix, B-v to B-ix, or C-v to C-ix and said second segment is any one of items D-v to D-ix, E-v to E-ix, or F-v to F-ix.

12) The protein according to item 11, wherein said first segment is any one of items A-v to A-ix and said second segment is any one of items D-v to D-ix, respectively; or said first segment is any one of items B-v to B-ix and said second segment is any one of items E-v to E-ix, respectively; or said first segment is any one of items C-v to C-ix and said second segment is any one of items F-v to F-ix, respectively.

13) The protein according to item 11 or 12, wherein said first segment is any one of items A-v to A-vi, B-v to B-vi, or C-v to C-vi and said second segment is any one of items D-v to D-vi, E-v to E-vi, or F-v to F-vi, respectively; and/or
wherein said first segment is any one of items A-viii to A-ix, B-viii to B-ix, or C-viii to C-ix and said second segment is any one of items D-viii to D-ix, E-viii to E-ix, or F-viii to F-ix, respectively.

14) The protein according to any one of item 4 to 13, wherein the cytotoxic activity of said protein is such that said protein and a comparative protein of the amino acid sequence of the SEQ ID NO: 1 produces spots free of viable bacteria of sensitive *Klebsiella quasipneumoniae* subsp. *similipneumoniae* SB30 (DSM 28212) at least of the same diameter 16 hours after spotting 20 microliters of each of a solution of said protein and of the comparative protein onto a lawn of the sensitive Klebsiella strain on an agar plate and subsequent incubation of the agar plate at 37°C, wherein the concentration of said protein in the solution is at most 5 times that of the solution of the respective comparative protein.

15) The protein according to any one of items 1 to 14, comprising or consisting of an amino acid sequence comprising or consisting of

(a) the amino acid sequence of
  (a-i) SEQ ID NO: 1 (KpneM),
  (a-ii) SEQ ID NO: 2 (KvarM),
  (a-iii) SEQ ID NO: 3 (KpneM2),
  (a-iv) SEQ ID NO: 4 (KaerM),
  (a-v) SEQ ID NO: 5 (KpneA),
  (a-vi) SEQ ID NO: 6 (KaerA),
  (a-vii) SEQ ID NO: 7 (Koxy),
  (a-viii) SEQ ID NO: 8 (Kpnela), or
  (a-ix) SEQ ID NO: 9 (Kvarla);

or

(b) an amino acid sequence

(b-i) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 1,

(b-ii) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 2,

(b-iii) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 3,

(b-iv) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 4,

(b-v) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 5,

(b-vi) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 6,

(b-vii) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 7,

(b-viii) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 8, or

(b-ix) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 9;

or

(c) an amino acid sequence

(c-i) having from 1 to 80 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 1,

(c-ii) having from 1 to 80 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 2,

(c-iii) having from 1 to 80 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 3,

(c-iv) having from 1 to 80 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 4,

(c-v) having from 1 to 110 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 5,

(c-vi) having from 1 to 110 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 6,

(c-vii) having from 1 to 130 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 7,

(c-viii) having from 1 to 130 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 8, or

(c-ix) having from 1 to 120 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 9.

16) The protein according to item 1 or 15, wherein

(a) the amino acid sequence of said protein is that of

(a-i) SEQ ID NO: 1 (KpneM),

(a-ii) SEQ ID NO: 2 (KvarM),

(a-iii) SEQ ID NO: 3 (KpneM2),

(a-iv) SEQ ID NO: 4 (KaerM),

(a-v) SEQ ID NO: 5 (KpneA),

(a-vi) SEQ ID NO: 6 (KaerA),

(a-vii) of SEQ ID NO: 7 (Koxy),

(a-viii) of SEQ ID NO: 8 (Kpnela), or

(a-ix) SEQ ID NO: 9 (Kvarla);

or

(b) the amino acid sequence of said protein has

(b-i) at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 1,

(b-ii) at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 2,

(b-iii) at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 3,

(b-iv) at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 4,

(b-v) at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 5,

(b-vi) at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 6,

(b-vii) at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 7,

(b-viii) at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 8, or

(b-ix) at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 9;

or

(c) the amino acid sequence of said protein has

(c-i) from 1 to 80 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 1,

(c-ii) from 1 to 80 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 2,

(c-iii) from 1 to 80 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 3,

(c-iv) from 1 to 80 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 4,

(c-v) from 1 to 110 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 5,

(c-vi) from 1 to 110 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 6,

(c-vii) from 1 to 130 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 7,

(c-viii) from 1 to 130 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 8, or

(c-ix) from 1 to 120 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 9.

17) The protein according to any one of items 15 or 16, wherein in item (b) any one of the sequence identities is at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, even more preferably at least 95%, and most preferably at least 97%; and/or
wherein in item (c) the number of said amino acid substitutions, additions, insertions and/or deletions is from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10, and most preferably at least 1 to 5 compared to any one of said amino acid sequences.

18) The protein according to any one of item 15, 16, or 17, wherein the cytotoxic activity of said protein of any one of items (b-i) to (b-ix) or (c-i) to (c-ix) is such that said protein and a comparative protein of the amino acid sequence of the SEQ ID NO of said item (b-i) to (b-ix) or (c-i) to (c-ix), respectively, produce spots free of viable bacteria of sensitive *Klebsiella quasipneumoniae* subsp. *similipneumoniae* SB30 (DSM 28212) at least of the same diameter 16 hours after spotting 20 microliters of each of a solution of said protein and of the comparative protein onto a lawn of the sensitive *Klebsiella* strain on an agar plate and subsequent incubation of the agar plate at 37°C, wherein the concentration of said protein in the solution is at most 5 times that of the solution of the respective comparative protein.

19) The protein according to any one of items 1 to 17, wherein said protein has a cell wall biosynthesis-inhibiting activity, whereby the protein is capable of degrading undecaprenyl phosphate-linked peptidoglycan precursors.

20) The protein according to any one of items 2 to 17, wherein the first segment comprises a translocation and a receptor-binding domain.

21) The protein according to any one of items 1 to 20, having bactericidal or bacteriostatic activity against *Klebsiella pneumoniae*, *Klebsiella oxytoca*, *Klebsiella granulomatis*, *Klebsiella quasipneumoniae*, *Klebsiella aerogenes*, and/or *Klebsiella variicola*.

22) A protein having cytotoxic activity against *Klebsiella*, optionally a protein according to any one of items 1 to 21, said protein comprising an amino acid sequence selected from the

group consisting of SEQ ID NO: 22 to 24, wherein each X stands for any one of the 20 standard amino acid residues or for an absent amino acid residue and J stands for either L (leucine) or I (isoleucine).

23) The protein according to any one of items 1 to 22, wherein said protein is any one defined with respect to SEQ ID NOs: 1-4, 7, 8 or 22 as a reference sequence, preferable any one defined with respect to SEQ ID NOs: 1-4 or 22 as a reference sequence.

24) The protein according to item 1 or 23, wherein the cytotoxic activity of said protein is such that said protein and a comparative protein of the amino acid sequence of SEQ ID NO: 1 produce spots free of viable bacteria of sensitive *Klebsiella quasipneumoniae* subsp. *similipneumoniae* SB30 (DSM 28212) at least of the same diameter 16 hours after spotting 20 microliters of each of a solution of said protein and of the comparative protein onto a lawn of the sensitive *Klebsiella* strain on an agar plate and subsequent incubation of the agar plate at 37°C, wherein the concentration of the protein in the solution is at most 5 times that of the solution of the comparative protein.

25) A composition comprising one or more proteins as defined in any one of items 1 to 24.

26) The composition according to item 25, wherein the cytotoxic activity of said composition is such that said composition and a comparative composition containing the comparative protein of the amino acid sequence of SEQ ID NO: 1 produce spots free of viable bacteria of sensitive *Klebsiella quasipneumoniae* subsp. *similipneumoniae* SB30 (DSM 28212) at least of the same diameter 16 hours after spotting 20 microliters of each of a solution of said composition and a comparative solution of the comparative composition onto a lawn of the sensitive *Klebsiella* strain on an agar plate and subsequent incubation of the agar plate at 37°C, wherein the protein concentration in the solution of said composition is at most 5 times that of the protein concentration in the comparative solution.

27) The composition according to item 25 or 26, further comprising at least one colicin and/or at least one salmocin.

28) The composition according to any one of items 25 to 27, wherein the composition is a pharmaceutical composition.

29) The composition according to any one of items 25 to 28, wherein said composition is a plant material or extract thereof, wherein the plant material is a material from a plant having expressed said one or more proteins, preferably an edible plant having expressed said one or more proteins.

30) The composition according to item 29, wherein said plant material is a material from a plant selected from the group consisting spinach, chard, beetroot, carrot, sugar beet, leafy beet, amaranth, Nicotiana, preferably Nicotiana benthamiana, and/or said plant material is one or more leaves, roots, tubers, or seeds, or a crushed, milled or comminuted product of said leaves, roots, tubers, or seeds.

31) The composition according to any one of items 25 to 30, wherein said composition is an aqueous solution containing said protein in dispersed form, preferably in dissolved form.

32) The protein or composition according to any one of items 1 to 31 for use in therapy, preferably for use in a method of treating infection of a subject with *Klebsiella*, such as *Klebsiella pneumoniae*, *Klebsiella oxytoca*, *Klebsiella quasipneumoniae*, *Klebsiella aerogenes*, and/or *Klebsiella variicola*.

33) The protein for the use according to item 32, wherein said *Klebsiella* is antibiotic-resistant such as carbapenem-resistant.

34) The protein for the use according to item 32 or 33, wherein said protein is any one defined with respect to SEQ ID NOs: 1-4, 7, 8 or 22 as a reference sequence, preferable any one defined with respect to SEQ ID NOs: 1-4 or 22 as a reference sequence.

35) A method of preventing or reducing infection or contamination of an object with one or more *Klebsiella species*, comprising contacting said object with a protein as defined in any one of items 1 to 24 or a composition as defined in any one of items 25 to 31.

36) A method of treating infection with *Klebsiella* of a subject in need thereof, comprising administering to said subject a protein as defined in any one of items 1 to 24 or a composition as defined in any one of items 25 to 31.

37) The method according to item 35 or 36, wherein said *Klebsiella* includes *Klebsiella pneumoniae*, *Klebsiella oxytoca*, *Klebsiella quasipneumoniae*, *Klebsiella aerogenes*, and/or *Klebsiella variicola*.

38) A process of producing a composition comprising a protein as defined in any one of items 1 to 24, said process comprising the following steps:

(i) expressing said protein in a plant, preferably an edible plant or *Nicotiana*,

(ii) harvesting plant material containing expressed protein from said plant,

(iii) extracting said protein from said plant material using an aqueous buffer to obtain a composition containing said protein,

(iv) optionally removing undesired contaminants from said composition.

39) The composition according to any one of items 25 to 31, wherein said one or more proteins is/are formulated for oral delivery to the small or large intestine.

40) Oral formulation comprising the protein according to any one of items 1 to 24 or the composition according to any one of items 25 to 31, said formulation being capable of protecting the protein from gastric conditions and capable of releasing the protein in the small or large intestine.

41) A nucleic acid molecule encoding the protein as defined in any one of items 1 to 24.

42) A nucleic acid molecule or nucleic acid construct encoding the protein as defined in any one of items 1 to 24, preferably the protein as defined in any one of items 15 to 23, said nucleic acid molecule or nucleic acid construct comprising a transcription promoter that is preferably active in plant cells and a nucleotide sequence encoding said protein for expressing said nucleotide sequence in cells, preferably in plant cells, under the control of said promoter.

43) A nucleic acid molecule or nucleic acid construct encoding the protein as defined in any one of items 1 to 24, preferably the protein as defined in any one of items 15 to 23, said nucleic acid molecule or nucleic acid construct is or encodes a viral (DNA or RNA) replicon comprising a nucleotide sequence encoding said protein for expressing said nucleotide sequence in cells, preferably in plant cells; said replicon may contain a subgenomic promoter for expressing said nucleotide sequence in plant cells or cells of a plant under the control of said subgenomic promoter.

44) A plant, plant tissue, or plant cell, comprising a protein as defined in any one of items 1 to 24.

45) A plant, plant tissue, or plant cell, comprising a nucleic acid molecule or nucleic acid construct as defined in item 42 or 43.

## BRIEF DESCRIPTION OF THE DRAWINGS

**Figure 1** shows schematically T-DNA regions of TMV-based vectors for the expression of klebicins. RB: right T-DNA border, Act2: *Arabidopsis thaliana* actin promoter; RdRp: RNA-dependent RNA polymerase; 3'NTR: 3' non-translated region; T: nos terminator; LB: left T-DNA border. KpneM-cat1: coding sequence of the klebicin KpneM (*K. pneumoniae* EWD35590.1) with

the first intron of catalase gene (cat-1) from *Ricinus communis*; KpneM2: coding sequence of the klebicin KpneM2 (*Klebsiella sp.* WP_047066220); KvarM: coding sequence of the klebicin KvarM (*K. variicola* CTQ17225.1); KaerM: coding sequence of the klebicin KaerM (*K. aerogenes* WP_015367360.1); KpneA: coding sequence of the klebicin KpneA (*K. pneumoniae* SAV78255.1); KaerA: coding sequence of the klebicin KaerA (*K. aerogenes* WP_063414841.1); KoxyY: coding sequence of the klebicin KoxyY (*K. oxytoca* WP_024273778); Kvarla: coding sequence of the klebicin Kvarla (*K. variicola* KDL88409); Kpnela: coding sequence of the klebicin Kpnela (*K. pneumoniae* BAS34675).

**Figure 2** shows SDS-PAGE analysis of the expression of klebicins in *N. benthamiana* leaves. Plant material (50 mg) was harvested at 5 or 7 days post spraying (dps) (pooled samples of three leaves – klebicin KaerA at 4 dps, all remaining klebicins at 5 dps), ground in liquid nitrogen, extracted with 50 mM Tris-HCl, 300 mM NaCl 15 mM sodium acetate, 3 mM DTT (pH 7.5), and denatured at 98°C for 10 min. Solutions containing 5 µg of protein were resolved in 12% polyacrylamide gel for Coomassie staining. M – PageRuler Prestained protein ladder (ThermoFisher Scientific Baltics), WT – crude extract of non - sprayed *N. benthamiana* leaves, Kvarla, Kpnela, KpneA, KaerA, KoxyY, KpneM, KpneM2, KvarA, KaerM – extracts of *N. benthamiana* leaves, sprayed with klebicin expression constructs. Bands corresponding to recombinant klebicins are marked by asterisks.

**Figure 3** illustrates the purification of klebicins from *Nicotiana benthamiana* leaf biomass. A, C, E, G, I, K – purification schemes of KpneM (A), KpneM2 (C), KvarM (E), KpneA (G), KaerA (I) and Kvarla (K). B, D, F, H, J, L - SDS-PAGE analysis of protein samples taken from different KpneM (B), KpneM2 (D), KvarM (F), KpneA (H), KaerA (J) and Kvarla (L) purification steps. Solutions containing 5 µg of protein were resolved in 12% SDS-PAGE gel for Coomassie staining. B - lane 1 and 7 - PageRuler™ Prestained protein ladder, lane 2 – crude extract, lane 3 – total soluble proteins loaded on Phenyl sepharose , lane 4 – flow through Phenyl sepharose, lane 5 – KpneM eluate (after Phenyl sepharose), lane 6 – impurities eluate (after Phenyl sepharose) , lane 8 – proteins loaded on Q sepharose, lane 9 - KpneM flow through Q sepharose, lane 10 - impurities eluate (after Q sepharose); D - lane 1 and 6 - PageRuler™ Prestained protein ladder, lane 2 – total soluble proteins loaded on Phenyl sepharose , lane 3 – flow through Phenyl sepharose, lane 4 – KpneM2 eluate (after Phenyl sepharose), lane 5 – impurities eluate (after Phenyl sepharose), lane 7 – proteins loaded on Q sepharose, lane 8 – KpneM2 flow through Q sepharose, lane 9 - impurities eluate (after Q sepharose); F - lane 1 and 7 - PageRuler™ Prestained protein ladder, lane 2 – crude extract, lane 3 - total soluble proteins loaded on Phenyl sepharose, lane 4 – flow through Phenyl sepharose, lane 5 – KvarM eluate (after Phenyl sepharose), lane 6 – impurities eluate (after Phenyl sepharose), lane 8 – proteins loaded on Q sepharose, lane 9 – KvarM flow through Q sepharose, lane 10 - impurities eluate (after Q

sepharose); H - lane 1 and 7 - PageRuler™ Prestained protein ladder, lane 2 – crude extract, lane 3 – total soluble proteins loaded on Phenyl sepharose, lane 4 – flow through Phenyl sepharose, lane 5 – KpneA eluate (after Phenyl sepharose), lane 6 – impurities eluate (after Phenyl sepharose), lane 8 – proteins loaded on SP sepharose , lane 9 – flow through SP sepharose, lane 10 - KpneA eluate (after SP sepharose); J – lane 1 and 6 - PageRuler™ Prestained protein ladder, lane 2 – crude extract, lane 3 – total soluble proteins loaded on SP sepharose, lane 4 – flow through SP sepharose, lane 5 – KaerA eluate (after SP sepharose), lane 7 – proteins loaded on Q sepharose, lane 8 – KaerA flow through Q sepharose. L – lane 1 and 7 - PageRuler™ Prestained protein ladder, lane 2 – crude extract, lane 3 – total soluble proteins loaded on Phenyl sepharose , lane 4 – flow through Phenyl sepharose, lane 5 – Kvarla eluate (after Phenyl sepharose), lane 6 – impurities eluate (after Phenyl sepharose), lane 8 – proteins loaded on SP sepharose, lane 9 – flow through SP sepharose, lane 10 - Kvarla eluate (after SP sepharose). The arrows mark recombinant proteins.

**Figure 4** shows purified klebicins on one gel. 0.5 µg of purified klebicins were resolved in 12% SDS-PAGE gel for Coomassie staining. M – PageRuler Unstained Protein Ladder (ThermoFisher Scientific Baltics).

**Figure 5** illustrates an evaluation of klebicins activity against *Klebsiella* strains in soft-agar overlay assay. Overnight cultures of bacterium were grown in CAA medium, equalized till OD595=1.0, diluted 100x with melted top CAA agar and poured on CAA agar plates. 20 µL drops of protein crude extract were spotted on 6 mm Whatman disks and Petri plates were incubated overnight at 30°C or 37° C.

**Figure 6** shows the sensitivity of clinical *Klebsiella* isolates to six plant-expressed klebicins. 100 clinical Klebsiella isolates (89 *K. pneumoniae* and 11 *K. oxytoca*) were tested in drop plate assay. The strains susceptible to each klebicin are grouped by the size of the inhibition zone.

**Figure 7** shows klebicin cytotoxicity assays in the liquid culture. Overnight Klebsiella cultures were diluted to OD600=0.3 in CAA medium, treated with 5 µg $mL^{-1}$ of either of klebicins and bacteria further incubated for 5 h with shaking (200 rpm). The antimicrobial activity of klebicins was evaluated by counting colony forming units in tested culture. Bars represent standard deviation.

**Figure 8** shows klebicins activity against biofilms. One day–old *K. quasipneumoniae, K. oxytoca, K. variicola, K. aerogenes* biofilms grown in CAA medium were treated with 5 µg $mL^{-1}$ of either of klebicins. The antimicrobial activity of klebicins was evaluated by counting colony forming units in tested culture. Bars represent standard deviation.

**Figure 9** shows Impact of Kvarla treatment to the survival of *Galleria mellonella* larvae after challenge with *K. quasipneumoniae* DSM 28212. *G. melonella* larvae were infected with

12000-32000 CFU of *K. pneumoniae* DSM 28212 and treated with 10 µg of Kvarla 2 hours after infection. Larvae were incubated in Petri dishes at 37 C° up to 68h. 20 larvae were used for each treatment point.

**Figure 10** shows the deduction of the consensus sequence as given in SEQ ID NO: 22-24. The one-letter code refers to the 20 standard amino acids, the „X" stands for insufficient conservation to deduce a consensus amino acid at the respective position and the „-" for omission of an amino acid during deduction of the consensus sequence due to low conservation. J represents either L (leucine) or I (isoleucine). Highly conserved amino acids are highlighted in black, moderately conserved amino acids are highlighted in grey. The software "Geneious Prime Clustal W" in standard settings was used for deduction of the consensus sequences. Fig. 10A shows deduction of SEQ ID NO: 22. Fig. 10B shows deduction of SEQ ID NO: 23. Fig. 10C shows deduction of SEQ ID NO: 24.

**Figure 11** shows the evaluation of the activity of klebicins KpneM, KpneM2, Kvarla, KpneA, KaerA and KvarM and their concentrations upon storage as lyophilized purified proteins. Klebicin activity against susceptible bacterium was evaluated in liquid cultures or by radial diffusion assay. Antimicrobial activity was expressed as CFU/mL $\Delta\log_{10}$ when activity was evaluated in liquid cultures or as specific activity units (AU) if radial diffusion assay was used for the evaluation. Klebicin concentration was determined by the Bradford assay. Data are the mean ± SD of three independent experiments. (**A**) Activity of klebicins KpneM, KpneM2, Kvarla upon storage at -20°C. (**B**) Activity of klebicins KpneM, KpneM2, Kvarla upon the storage at 5°C. (**C**) Activity of klebicins KpneM, KpneM2, Kvarla upon the storage at room temperature. (**D**) Activity of klebicins KpneA, KaerA and KvarM upon the storage at -20°C. (**E**) Activity of klebicins KpneA, KaerA and KvarM upon the storage at 5 °C. (**F**) Activity of klebicins KpneA, KaerA and KvarM upon the storage at room temperature. (**G**) Trendlines of klebicin concentration upon the storage at -20°C. (**H**) Trendlines of klebicin concentration upon the storage at 5°C. (**I**) Trendlines of klebicin concentration upon the storage at room temperature.

**Figure 12** shows the residual activity of Kvarla and Eudragit S100-coated Kvarla after *in vitro* gastric digestion in soft agar overlay assay and assessment of the digestion of the Kvarla by pepsin into fragments by SDS-PAGE. (**A**) The evaluation of residual activity of Kvarla and Eudragit S100-coated Kvarla in soft agar overlay assay after simulated gastric digestion *in vitro*. Protein samples were digested by pepsin (pepsin:protein ratio 1:40) for 0.5, 5, 10, 20, 30 and 60 min. Dilutions of all samples by ratio 1:2 were made in distilled water and 5 µL aliquots of diluted samples were dropped on MHA plates with *K. quasipneumoniae* DSM28212 lawn. (**B**) Tricine SDS page assay of Kvarla digestion. Coomassie staining was used to visualize protein decomposition and estimate the MW of peptide products. The presence and/or absence of pepsin and Kvarla are indicated. Times are indicated in minutes and correspond to those in (**A**).

**Figure 13** shows the standard curve for the detection of *K. quasipneumoniae* obtained by real time-PCR, based on *khe* gene amplification.

**Figure 14** shows real time-PCR results in *K. quasipneumoniae* colonized mice faecal samples before and after klebicin treatment. Faecal samples of 3 mice were used for each experimental point. **18d**: faecal samples collected 18$^{th}$ day of experiment, before the start of klebicin treatment; **22d**: faecal samples collected 22$^{nd}$ day of experiment, next day after last klebicin gavage.

DETAILED DESCRIPTION OF THE INVENTION

The inventors of the present invention have identified proteins having bactericidal or bacteriostatic activity against *Klebsiella.* Such proteins are referred to herein as "klebicins". The protein or klebicin of the invention preferably has a lipid II-cleaving activity or a capability of forming pores in a bacterial cell membrane.

The protein or klebicin of the invention generally comprises at least two amino acid sequence segments (sometimes briefly referred to herein as "segments"). Herein, an amino acid sequence segment refers to a plurality of contiguous amino acid residues in the primary structure of a protein or polypeptide, wherein the protein or polypeptide has a larger number of amino acid residues in its primary structure than the segment. The protein of the invention generally comprises or consists of a first segment and a second segment. The first segment generally provides the protein with the capability of binding to components of *Klebsiella* cells (such as binding to a receptor) and/or it provides the protein with the capability of being introduced into or being taken up by *Klebsiella* cells (cell translocation). The second segment may have lipid II-cleaving activity or a pore-forming capability in a bacterial cell membrane. The second segment therefore provides the protein with its cytotoxic activity. In one embodiment, the first segment is on the N-terminal side in the primary structure of said protein and the second segment is on the C-terminal side or vice versa, wherein the former is preferred. Accordingly, the inventive protein may comprise or consist of an N-terminal first segment and a C-terminal second segment. In another embodiment, the second segment is on the N-terminal side in the primary structure of said protein and the first segment is on the C-terminal side. Accordingly, the inventive protein may comprise or consist of an N-terminal second segment and a C-terminal first segment.

First segment of the protein of the invention

The protein of the invention may comprise a first segment that comprises the amino acid sequence of any one of items (A-i) to (A-ix) defined above. The amino acid sequences of SEQ ID NO: 1 to 9 recited in these items are amino acid sequences of the klebicins identified by the inventors. Preferably, the amino acid sequence of the first segment is the amino acid sequence of items (A-i) to (A-ix) defined above.

However, the invention is not limited to klebicins having first segments of the specific klebicins identified by the inventors. Instead of the amino acid sequences of items (A-i) to (A-ix), the first segment may comprise an amino acid sequence of any one of items (B-i) to (B-ix), respectively, defined above. The wording "wherein the first segment comprises an amino acid sequence having at least 70% sequence identity to the amino acid sequence from amino acid residue x to y of SEQ ID NO: Z" (wherein x and y stand for the start and end positions indicated and Z stands for the number of the SEQ ID) means that the amino acid sequence of the first segment has preferably at least the same number of amino acid residues as the sequence from amino acid residue x to y of SEQ ID NO: Z and has at least the indicated sequence identity over the entire length from residue x to y of SEQ ID NO: Z. This principle applies to all items (B-i) to (B-ix) and to other sequence identities defined herein. Herein, the determination of sequence identities is done using Clustal Omega (CLUSTAL O 1.2.4) based on standard parameters. Preferably, the amino acid sequence of the first segment is that of any one of items (B-i) to (B-ix). The wording "wherein the amino acid sequence of the first segment has at least 70% sequence identity to the amino acid sequence from amino acid residue x to y of SEQ ID NO: Z" means that the amino acid sequence of the first segment has at least the same number of amino acid residues as the sequence from amino acid residue x to y of SEQ ID NO: Z and has the indicated sequence identity over the entire length from residue x to y of SEQ ID NO: Z. This applies to all items (B-i) to (B-ix) and to corresponding sequence identities defined herein.

In another embodiment, the first segment comprises an amino acid sequence of any one of items (C-i) to (C-ix) as defined above. Preferably, the amino acid sequence of the first segment is as defined in any one of items (C-i) to (C-ix) as defined above. The definitions of items (C-i) to (C-ix) mean that the amino acid sequence is that of the indicated amino acid residue range of the indicated SEQ ID NO except for the indicated number of substitutions, additions, insertions and/or deletions.

Where the protein is defined herein by a number or numerical range of amino acid substitutions, additions, insertions and/or deletions, these amino acid substitutions, additions, insertions or deletions may be combined, but the given number or numerical range refers to the sum of all amino acid substitutions, additions, insertions and deletions. Among amino acid substitutions, additions, insertions and deletions, amino acid substitutions, additions, and deletions are preferred. The term "insertion" relates to insertions within the amino acid sequence of a reference sequence, i.e. excluding additions at the C- or N-terminal end. The term "addition" means additions at the C- or N-terminal end of the amino acid sequence of a reference sequence. A deletion may be a deletion of a terminal or an internal amino acid residue of a reference sequence. The term "reference sequence" is used herein to refer to an amino acid sequence of the sequence listing based on which an amino acid sequence of a

protein of the invention is defined. For example, in item (A-i) the reference sequence is SEQ ID NO: 1.

In items (B), any one of the sequence identities may be at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, even more preferably at least 95%, and most preferably at least 97%. In items (C), the number of said amino acid substitutions, additions, insertions and/or deletions is from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10, and most preferably from 1 to 5 compared to any one of said amino acid sequences.

Accordingly, the following items (i) to (ix) of each of items (B) and (C) define preferred embodiment of the first segment of the protein of the invention. The first segment of the protein of the invention preferably comprises any of the following amino acid sequences:

(B-i) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 128 of SEQ ID NO: 1,

(B-ii) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 127 of SEQ ID NO: 2,

(B-iii) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 123 of SEQ ID NO: 3,

(B-iv) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 118 of SEQ ID NO: 4,

(B-v) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 170 of SEQ ID NO: 5,

(B-vi) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 172 of SEQ ID NO: 6,

(B-vii) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 255 of SEQ ID NO: 7,

(B-viii) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 288 of SEQ ID NO: 8, or

(B-ix) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 236 of SEQ ID NO: 9;

or

(C-i) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 1 to 128 of SEQ ID NO: 1,

(C-ii) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 1 to 127 of SEQ ID NO: 2,

(C-iii) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 1 to 123 of SEQ ID NO: 3,

(C-iv) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 1 to 118 of SEQ ID NO: 4,

(C-v) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 1 to 170 of SEQ ID NO: 5,

(C-vi) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 1 to 172 of SEQ ID NO: 6,

(C-vii) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 1 to 255 of SEQ ID NO: 7,

(C-viii) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 1 to 288 of SEQ ID NO: 8, or

(C-ix) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 1 to 236 of SEQ ID NO: 9.

In another embodiment, the amino acid sequence of the first segment of the protein of the invention preferably has:

(B-i) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 128 of SEQ ID NO: 1,

(B-ii) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 127 of SEQ ID NO: 2,
(B-iii) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 123 of SEQ ID NO: 3,
(B-iv) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 118 of SEQ ID NO: 4,
(B-v) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 170 of SEQ ID NO: 5,
(B-vi) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 172 of SEQ ID NO: 6,
(B-vii) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 255 of SEQ ID NO: 7,
(B-viii) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 288 of SEQ ID NO: 8, or
(B-ix) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the amino acid sequence from amino acid residue 1 to 236 of SEQ ID NO: 9;
or
(C-i) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 128 of SEQ ID NO: 1,
(C-ii) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 127 of SEQ ID NO: 2,
(C-iii) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 123 of SEQ ID NO: 3,
(C-iv) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 118 of SEQ ID NO: 4,

(C-v) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 170 of SEQ ID NO: 5,
(C-vi) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 172 of SEQ ID NO: 6,
(C-vii) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 255 of SEQ ID NO: 7,
(C-viii) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 288 of SEQ ID NO: 8, or
(C-ix) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 236 of SEQ ID NO: 9.

Second segment of the protein of the invention

Since the protein (klebicin) of the invention has both a first and a second segment, the protein may be defined by both its first and second segment. Thus, the protein may be defined by a combination of any one of the above first segments with a second segment as defined herein.

The second segment of the protein of the invention provides the protein with its bactericidal or bacteriostatic activity against *Klebsiella*. The second segment preferably has a lipid II-cleaving activity or a pore-forming capability in a bacterial cell membrane. Such activities are known from other bacteriostatic or bactericidal bacterial proteins such as *E. coli* colicins.

The second segment may comprise or consist of the amino acid sequence of any one of items (D-i) to (D-ix) defined above. Preferably, the amino acid sequence of the first segment is the amino acid sequence of items (D-i) to (D-ix) defined above. Where a definition of a first segment is combined with that of a second segment, first and second segments defined based on same SEQ ID NO as reference sequence are preferably combined. For example, a protein may be defined by having a first segment based on item (A-ii) and a second segment based on item (D-ii).

However, the invention is not limited to klebicins having second segments of the specific klebicins identified by the inventors. Instead of the amino acid sequences of items (D-i) to (D-ix), the second segment may comprise or consist of an amino acid sequence of any one of items (E-i) to (E-ix). Preferably, the amino acid sequence of the second segment is that of any one of items (E-i) to (E-ix).

In another embodiment, the second segment comprises or consists of an amino acid sequence of any one of items (F-i) to (F-ix) as defined above. Preferably, the amino acid sequence of the first segment is as defined in any one of items (F-i) to (F-ix) as defined above.

The following items (i) to (ix) of each of items (E) and (F) define preferred second segments. In one embodiment, the second segment of the protein of the invention preferably comprises any of the following amino acid sequences:

(E-i) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 129 to 278 of SEQ ID NO: 1,

(E-ii) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 128 to 276 of SEQ ID NO: 2,

(E-iii) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 124 to 272 of SEQ ID NO: 3,

(E-iv) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 119 to 266 of SEQ ID NO: 4,

(E-v) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 171 to 377 of SEQ ID NO: 5,

(E-vi) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 173 to 379 of SEQ ID NO: 6,

(E-vii) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 256 to 452 of SEQ ID NO: 7,

(E-viii) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 289 to 466 of SEQ ID NO: 8, or

(E-ix) having at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 237 to 414 of SEQ ID NO: 9;

or

(F-i) having from 1 to 20, preferably from 1 to 15, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 129 to 278 of SEQ ID NO: 1,

(F-ii) having from 1 to 20, preferably from 1 to 15, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 128 to 276 of SEQ ID NO: 2,
(F-iii) having from 1 to 20, preferably from 1 to 15, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 124 to 272 of SEQ ID NO: 3,
(F-iv) having from 1 to 20, preferably from 1 to 15, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 119 to 266 SEQ ID NO: 4,
(F-v) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 171 to 377 of SEQ ID NO: 5,
(F-vi) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 173 to 379 SEQ ID NO: 6,
(F-vii) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 256 to 452 of SEQ ID NO: 7,
(F-viii) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 289 to 466 of SEQ ID NO: 8, or
(F-ix) having from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 237 to 414 of SEQ ID NO: 9.

In another embodiment, the amino acid sequence of the second segment of the protein of the invention preferably has:
(E-i) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 129 to 278 of SEQ ID NO: 1,
(E-ii) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 128 to 276 of SEQ ID NO: 2,
(E-iii) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 124 to 272 of SEQ ID NO: 3,
(E-iv) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 119 to 266 of SEQ ID NO: 4,

(E-v) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 171 to 377 of SEQ ID NO: 5,
(E-vi) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 173 to 379 of SEQ ID NO: 6,
(E-vii) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 256 to 452 of SEQ ID NO: 7,
(E-viii) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 289 to 466 of SEQ ID NO: 8,
(E-ix) at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity to the segment from amino acid residue 237 to 414 of SEQ ID NO: 9;
or
(F-i) from 1 to 20, preferably from 1 to 15, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 129 to 278 of SEQ ID NO: 1,
(F-ii) from 1 to 20, preferably from 1 to 15, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 128 to 276 of SEQ ID NO: 2,
(F-iii) from 1 to 20, preferably from 1 to 15, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 124 to 272 of SEQ ID NO: 3,
(F-iv) from 1 to 20, preferably from 1 to 15, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 119 to 266 SEQ ID NO: 4,
(F-v) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 171 to 377 of SEQ ID NO: 5,
(F-vi) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 173 to 379 SEQ ID NO: 6,
(F-vii) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 256 to 452 of SEQ ID NO: 7,

(F-viii) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 289 to 466 of SEQ ID NO: 8,
(F-ix) from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid 237 to 414 of SEQ ID NO: 9.

In the invention, any first segment mentioned above may be combined with any second segment to a protein of the invention. In one embodiment, the protein of the invention comprises a first segment of any one of items (i)-(iv) and a second segment of any one of items (i)-(iv), irrespective of whether the first segment belongs to item (A), (B) or (C) and whether the second segment belongs to items (D), (E), or (F). However, in one embodiment, the protein of the invention comprises a first segment of any one of items (C-i)-(C-iv) and a second segment of any one of items (F-i)-(F-iv).

In another embodiment, the protein of the invention comprises a first segment of any one of items (v)-(ix) and a second segment of any one of items (v)-(ix), irrespective of whether the first segment belongs to item (A), (B) or (C) and whether the second segment belongs to items (D), (E), or (F). However, in one embodiment, the protein of the invention comprises a first segment of any one of items (C-v)-(C-ix) and a second segment of any one of items (F-v)-(F-ix).

Other embodiments are as follows:

in one embodiment, the protein of the invention may have a first segment that comprises or consists of any of the amino acid sequences of items (A-i) to (A-iv), (B-i) to (B-iv), or (C-i) to (C-iv), in a most generic of preferred embodiments, and a second segment that comprises or consists of any one of the amino acid sequences of items (D-i) to (D-iv), (E-i) to (E-iv), or (F-i) to (F-iv). Preferably, the first segment comprises or consists of any one of the amino acid sequences of items (A-i) to (A-iii), (B-i) to (B-iii), or (C-i) to (C-iii) and the second segment comprises any one of the amino acid sequences of (D-i) to (D-iii), (E-i) to (E-iii), or (F-i) to (F-iii). More preferably, the first segment comprises any one of the amino acid sequences of (A-ii), (B-ii) or (C-ii), and the second segment comprises any one of the amino acid sequence of (D-ii), (E-ii) or (F-ii). Even more preferably, the first segment comprises the amino acid sequence of (A-ii) and the second segment comprises the amino acid sequence of (D-ii).

In another embodiment, the first segment comprises any one of the amino acid sequences of items (A-v) to (A-ix), (B-v) to (B-ix), or (C-v) to (C-ix), and the second segment comprises any one of the amino acid sequences of (D-v) to (D-ix), (E-v) to (E-ix), or (F-v) to (F-ix). In a preferred embodiment, the first segment comprises any one of the amino acid sequences of items (A-v), (A-vi), (B-v), (B-vi), (C-v) and (C-vi), and the second segment comprises any one of the amino acid sequences of items (D-v), (D-vi), (E-v), (E-vi), (F-v) and

(F-vi). In a more preferred embodiment, the first segment comprises or consists of the amino acid sequence of item (A-v) or (A-vi) and the second segment comprises or consists of the amino acid sequence of item (D-v) or (D-vi). In another preferred embodiment, the first segment comprises or consists of the amino acid sequence of item (A-vii) and the second segment comprises or consists of the amino acid sequence of item (D-vii).

In a further alternative embodiment, the first segment comprises any one of the amino acid sequences of items (A-viii), (A-ix), (B-viii), (B-ix), (C-viii) or (C-ix), and the second segment comprises any one of the amino acid sequences of items (D-viii), (D-ix), (E-viii), (E-ix), (F-viii) or (F-ix). Preferably, the first segment comprises or consists of the amino acid sequence of item (A-viii) or (A-ix) and the second segment comprises or consists of the amino acid sequence of item (D-viii) or (D-ix).

Further preferred embodiments of the protein of the invention are as defined above in items (a-i) to (a-ix), (b-i) to (b-ix), and (c-i) to (c-ix). In items (b), the wording "amino acid sequence having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: Z" (wherein Z stands for the number of the SEQ ID NO) means that the amino acid sequence has at least the same number of amino acid residues as the sequence of SEQ ID NO: Z and has at least the indicated sequence identity over the entire length of SEQ ID NO: Z. This applies to all items (b-i) to (b-ix) and to corresponding sequence identities defined herein. The definitions of items (c-i) to (c-ix) mean that the amino acid sequence is that of the entire amino sequence of the indicated SEQ ID NO except for the indicated number of substitutions, additions, insertions or deletions.

Alternatively, the amino acid sequence of the inventive protein may be defined as follows:

(b') in one embodiment, the amino acid sequence of the protein of the invention is

(b'-i) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% identical to the amino acid sequence of SEQ ID NO: 1,

(b'-ii) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% identical to the amino acid sequence of SEQ ID NO: 2,

(b'-iii) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% identical to the amino acid sequence of SEQ ID NO: 3,

(b'-iv) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% identical to the amino acid sequence of SEQ ID NO: 4,

(b'-v) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% identical to the amino acid sequence of SEQ ID NO: 5,

(b'-vi) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% identical to the amino acid sequence of SEQ ID NO: 6,

(b'-vii) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% identical to the amino acid sequence of SEQ ID NO: 7,
(b'-viii) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% identical to the amino acid sequence of SEQ ID NO: 8, or
(b'-ix) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% identical to the amino acid sequence of SEQ ID NO: 9;

(b") in another embodiment, the amino acid sequence of the protein of the invention shares
(b"-i) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 1,
(b"-ii) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 2,
(b"-iii) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 3,
(b"-iv) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 4,
(b"-v) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 5,
(b"-vi) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 6,
(b"-vii) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 7,
(b"-viii)at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 8, or
(b"-ix) at least 70%, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 9.

Cytotoxic activity of the protein of the invention

The klebicin of the invention is capable of exerting a cytotoxic effect on bacteria of genus *Klebsiella*, such as *K. pneumoniae, K. granulomatis, K. oxytoca, K. aerogenes, K. quasipneumoniae, and K. variicola*. Preferably, the klebicin of the invention is active against *K. pneumoniae* and *K. oxytoca*. *K. pneumoniae* is most preferred. The cytotoxic effect may be a bacteriostatic or a bacteriocidal effect. Whether the protein has a cytotoxic effect can be tested experimentally, e.g. using the assay of Example 4. In one embodiment, said *Klebsiella* is antibiotic-resistant such as resistant to carbapenem, and said protein is a protein selected from items (i) to (iv), (vii) and (viii), preferably from (i) to (iv), of any of the embodiments defined above. The proteins selected from items (i) to (iv), (vii) and (viii) are proteins defined using SEQ ID NOs: 1 to 4, 7 or 8 as a reference sequence.

The protein of the invention may have a pore forming activity in cell membranes of *Klebsiella* cells. In the present invention, the proteins of items (a-v) to (a-ix) of SEQ ID NOs: 5 to 9 and the respective derivatives of items (b-v) to (b-ix) and (c-v) to (c-ix) have pore-forming activity. The pore-forming activity is envisaged to be due to the presence of the second amino acid sequence segment of these proteins.

Another class of proteins of the invention are envisaged to have a lipid II-cleaving activity, by analogy to the activity of *E. coli* colicin M. The inventors envisage that the proteins of this class have a peptidoglycanase activity that specifically cleaves the bond between the lipid moiety and the pyrophosphoryl group of the peptidoglycan lipid I and lipid II intermediates, located at the periplasmic side of the inner membrane, as determined for *E. coli* colicin M (Gross and Braun, Mol. Gen. Genet. 251 (1996) 388-396; Barreteau et al., Microbial Drug Resistance 18 (2012), 222-229). The released C55-polyisoprenol no longer translocates MurNAc-pentapeptide-GlcNAc across the cytoplasmic membrane. These klebicins are thus envisaged to exert toxicity against *Klebsiella* cells after they have been taken up across the outer cell wall into the periplasm. This property of the protein of the invention may be assayed according to the standard assay for colicin M activity described by El Ghachi *et al.*, J. Biol. Chem. 281 (2006) 22761-22772 using lipid I as the substrate. In the present invention, the proteins of items (a-i) to (a-iv) of SEQ ID NOs 1 to 4 and the respective derivatives of items (b-i) to (b-iv) and (c-i) to (c-iv) have peptidoglycanase or lipid II-cleaving activity. This activity is envisaged to be due to the presence of the second amino acid sequence segment of these proteins.

The cytotoxic activity of the protein of the invention is preferably such that said protein and a comparative protein of the amino acid sequence of the SEQ ID NO: 1 produces spots free of viable bacteria of sensitive *Klebsiella quasipneumoniae* subsp. *similipneumoniae* SB30 (DSM 28212) at least of the same diameter 16 hours after spotting 20 microliters of each of a solution of said protein and of the comparative protein onto a lawn of the sensitive Klebsiella strain on an agar plate and subsequent incubation of the agar plate at 37°C, wherein the concentration of said protein in the solution is at most 5 times that of the solution of the respective comparative protein. The solution is an aqueous solution. This test can be carried out as described in Reference Example 1. Protein concentrations are determined in terms of weight per volume as also described in Reference Example 1.

In one embodiment, the protein of the invention is one of any one of items (b-i) to (b-ix) or (c-i) to (c-ix) above and has a cytotoxic activity such that said protein and a comparative protein of the amino acid sequence of the reference sequence of the SEQ ID NO of said item (b-i) to (b-ix) or (c-i) to (c-ix), respectively, produce spots free of viable bacteria of the sensitive *Klebsiella* strain *quasipneumoniae* subsp. *similipneumoniae* SB30 (DSM 28212) at least of the same diameter 16 hours after spotting 20 microliters of each of a solution of said protein and of

the comparative protein onto a lawn of the sensitive *Klebsiella* strain on an agar plate and subsequent incubation of the agar plate at 37°C, wherein the concentration of said protein in the solution is at most 5 times that of the solution of the respective comparative protein. The solution is an aqueous solution. This test can be carried out as described in Reference Example 1. Protein concentrations are determined in terms of weight per volume as also described in Reference Example 1.

Consensus sequences of proteins of the invention

The protein of the invention comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 22 to 24, wherein each X stands for any one of the 20 standard amino acid residues or absence of an amino acid residue and J stands for either L (leucine) or I (isoleucine), preferably X stands for any one of the 20 standard amino acid residues and J for either L (leucine) or I (isoleucine). The 20 standard amino acid residues are: A (alanine), C (cysteine), D (aspartic acid), E (glutamic acid), F (phenylalanine), G (glycine), H (histidine), I (isoleucine), K (lysine), L (leucine), M (methionine), N (asparagine), P (proline), Q (glutamine), R (arginine), S (serine), T (threonine), V (valine), W (tryptophan) and Y (tyrosine).

In a preferred embodiment, the cytotoxic activity of said protein is such that said protein and a comparative protein of the amino acid sequence of SEQ ID NO: 1 produces spots free of viable bacteria of sensitive *Klebsiella quasipneumoniae* subsp. *similipneumoniae* SB30 (DSM 28212) at least of the same diameter 16 hours after spotting 20 microliters of each of a solution of said protein and of the comparative protein onto a lawn of the sensitive *Klebsiella* strain on an agar plate and subsequent incubation of the agar plate at 37°C, wherein the concentration of the protein in the solution is at most 5 times that of the solution of the comparative protein. The assay on the cytotoxic activity is carried out as described in Reference Example 1.

Proteins of the present invention with a similar mode-of-action against *Klebsiella* have conserved positions and/or amino acid sequence stretches in the amino acid sequence, preferably in the first and second segments. Conserved positions and/or amino acid sequence stretches have an increased likelihood of being relevant for the function of the protein of the invention. In the first amino acid sequence segment, the conserved positions or stretches generally relate to the functions of receptor binding and translocation. In the second amino acid sequence segment, the conserved positons or stretches generally relate to cytotoxicity against Klebsiella. The klebicins KpneM2, KvarM, KpneM and KaerM share a consensus sequence of SEQ ID NO: 22. The klebicins KaerA and KpneA share a consensus sequence of SEQ ID NO: 23. The klebicins Kpnela and Kvarla share a consensus sequence of SEQ ID NO: 24. Each X in SEQ ID NO: 22, 23 and 24 stands either for any one of the 20 standard amino acid residues or for no amino acid residue present at this position and J stands for either L (leucine) or I (isoleucine).

In particular, a klebicin with lipid II-cleaving activity (e.g. those defined with reference to SEQ ID NOs: 1-4) preferably has amino acid residues or amino acid sequence stretches corresponding to the following amino acid residues or amino acid sequence stretches: 1M, 2S/T, 3D/E, 4T, 5L/M, 7V, 9A, 22G, 24G, 50S, 91T, 97P, 132P, 138H, 139Y, 142G, 144G, 155G, 156L, 176G, 184F, 199L, 200G, 202I, 203T, 206TEGTL210 (SEQ ID NO: 25), 212I, 216G, 218W, 220YNGV223 (SEQ ID NO: 26), 225RAFNDTYD232 (SEQ ID NO: 27), 234N, 239R, 243A, 247T, 255G, 258Y, 260I, 262P, 262G, 271S, 272G;

wherein the number preceding a one-letter code of an amino acid residue indicates the position in SEQ ID NO: 22, and a number following a one-letter code of an amino acid residue indicates the position in SEQ ID NO: 22 of the preceding amino acid residue in case of a stretch of two or more amino acid residues. Two or more amino acid residues separated by "/" means that any one of the indicated residues is at the position corresponding to the indicated position in SEQ ID NO: 22. The wording "amino acid residues or amino acid sequence stretches corresponding to ... " means that the position in a given protein may differ from the position of SEQ ID NO: 22. However, the corresponding position can be determined by aligning the amino acid sequence of the protein with those of SEQ ID NOs: 1-4 and 22 (as shown in Fig. 10A) and determining the corresponding position by counting the residues of the amino acid sequence of the protein starting with the N-terminal first amino acid sequence.

A klebicin with pore forming activity defined with reference to SEQ ID NOs: 5 and 6 has preferably amino acid residues or amino acid sequence stretches corresponding to the following amino acid residues or amino acid sequence stretches: 1M, 4E, 9V, 11G, 13N, 18V, 20WGG22, 25GNGNNGGAG33 (SEQ ID NO: 28), 36G, 39G, 45G, 47T, 52L, 65P, 67N, 68P, 72GAPW75 (SEQ ID NO: 29), 80S, 82K, 84A, 90AN91, 94KP95, 97KFKANIQN104 (SEQ ID NO: 30), 106K, 111GSL113, 115SP116, 188V, 120KS121, 123SSGDVDTY130 (SEQ ID NO: 31), 132VSFGKEKYNV141 (SEQ ID NO: 32), 143YNRKKDSFT151 (SEQ NO ID: 33), 154YVDGGA159 (SEQ ID NO: 34), 161KPEHSMKDQAIAVV174 (SEQ ID NO: 35), 176LYLLNE181 (SEQ ID NO: 36), 186VI187, 189T, 193II194, 197SG198, 200T, 202SGKLG206 (SEQ ID NO: 37), 208KY209, 212LA213, 217A, 220I, 222NFQGKK227 (SEQ ID NO: 38), 229RSF231, 233DAM235, 237S, 244NP245, 247MKL249, 251QADK254 (SEQ ID NO: 39), 259NAL261, 263Q, 266LS267, 269LADRFKGL277 (SEQ ID NO: 40), 279AFTW282 (SEQ ID NO: 41), 284DRLLKA289 (SEQ ID NO: 42), 291KI292, 294DGVVTGVTTG303 (SEQ ID NO: 43), 305WQ306, 308LA309, 311EVEAMYLSGVAG322 (SEQ ID NO: 44), 324VALGI328 (SEQ ID NO: 45), 330T, 332MIS334, 337A, 341S, 343P, 346AV, 349ALTV352 (SEQ ID NO: 46), 354AVIGI358 (SEQ ID NO: 47), 360I, 362TSYI365 (SEQ ID NO: 48), 367AD368, 370AKALNNAV377 (SEQ ID NO: 49), 380LFK382;

wherein the number preceding a one-letter code of an amino acid residue indicates the position in SEQ ID NO: 23, and a number following a one-letter code of an amino acid residue indicates the position in SEQ ID NO: 23 of the preceding amino acid residue in case of a

stretch of two or more amino acid residues. Two or more amino acid residues separated by "/" means that any one of the indicated residues is at the position corresponding to the indicated position in SEQ ID NO: 23. The wording "amino acid residues or amino acid sequence stretches corresponding to ... " means that the position in a given protein may differ from the position of SEQ ID NO: 23. However, the corresponding position can be determined by aligning the amino acid sequence of the protein with those of SEQ ID NOs: 5-6 and 23 (as shown in Fig. 10B) and determining the corresponding position by counting the residues of the amino acid sequence of the protein starting with the N-terminal first amino acid sequence. SEQ ID NO: 23 is that given at the end of this specification.

A klebicin with pore forming activity defined with reference to SEQ ID NOs: 8 and 9 has preferably amino acid residues or amino acid sequence stretches corresponding to the following amino acid residues or amino acid sequence stretches:
1MPGFNYGGKGDGTNWSSERGTGPEPGGGSRGNGGDRDNSRGGAGNRGNWAGSGPLSA ALINDSIAEALEKQLPRNTVEATSTPAYKKMRAAFDALPLDKQPEARAQITKAWQSAHDAMPD1 20 (SEQ ID NO: 50), 121K/R,
122TTTTENVGGGKNGHNVTRSTPNWLKEKMKGLNQQVNNDLSGALAQHQKAEADARAKAE AAAKAK185 (SEQ ID NO: 51), 238A, 239E/A, 240AKAKAEAEAKAKAEA254 (SEQ ID NO: 52), 255A/E, 256AKAKAEA262 (SEQ ID NO: 53), 263E/A,
264AKAKAEAEAKAKAEAEAKAKAEADAVKDAVKFTADFYKEVFSVYGEKAEQLANLLATQAK GKNIRNIDDALKAYEKHKTNINKKINAQDRAAIAKALESVDVKEAAKNFAKFSKGLGYVGPTMD VVDLVLELRKAIKEDNWR405 (SEQ ID NO: 54), 406S/T,
407FFVKIEAIAISFGATQLAALAFASLLGAPVGLLGYALIMAGIGALVSDDVVDAANKIIGI466 (SEQ ID NO: 55);

wherein a number preceding a one-letter code of an amino acid residue indicate the position in SEQ ID NO: 24, and a number following a one-letter code of an amino acid residue indicates the position in SEQ ID NO: 24 of the preceding amino acid residue in case of a stretch of two or more amino acid residues. Two or more amino acid residues separated by "/" means that any one of the indicated residues is at the position corresponding to the indicated position in SEQ ID NO: 24. The wording "amino acid residues or amino acid sequence stretches corresponding to ... " means that the position in a given protein may differ from the position of SEQ ID NO: 24. However, the corresponding position can be determined by aligning the amino acid sequence of the protein with those of SEQ ID NOs: 8-9 and 24 (as shown in Fig. 10C) and determining the corresponding position by counting the residues of the amino acid sequence of the protein starting with the N-terminal first amino acid sequence.

The definitions given above with respect to the consensus sequences can be combined with the definitions given in the claims or the embodiments given in preceding sections.

I more detail, the definition of conserved residues may be combined with any definition of item (B-i) to (B-iv), (C-i) to (C-iv), (E-i) to (E-iv), (F-i) to (F-iv), (b-i) to (b-iv), and/or (c-i) to (c-

iv) above to define amino acid residue that should not be changed. Depending on the specific reference sequence used for defining the protein of the invention, the indications of the amino acid positions given above are exchanged by the corresponding positions of the respective reference sequence. Corresponding positions in the reference sequence can e.g. be derived from the alignment shown in Fig. 10A-C.

Klebicin compositions

The composition of the invention comprises one or more proteins (klebicins) of the invention as described above and optionally further components as the case requires such as a carrier. The composition may comprise one or more different proteins (klebicins) as defined herein, such as two, three or four different proteins (klebicins) as defined herein. "Different" means that the proteins differ in at least one amino acid residue. The composition may comprise two, three or more klebicins of the invention from the same class represented by any one of items (i) to (iv) above or of items (v) to (ix) above. Preferably, the composition contains at least two klebicins of the invention from different classes, such as at least one klebicin of the pore-forming type and at least one klebicin of the lipid II-cleaving type. The composition may further comprise one or more *E. coli* colicin or a derivative thereof e.g. as described in EP 3 097 783 A1, e.g. for concomitantly controlling pathogenic *E. coli* such as EHEC.

The invention also provides a composition comprising one or more proteins of the invention and one or more other bacteriocidal or bacteriostatic proteins. Such other bacteriocidal or bacteriostatic proteins may be *E. coli* colicins or *Salmonella* colicins (salmocins). *E. coli* colicin are known in the art and are described inter alia in EP3097783 A1. Salmocins are known and described in WO2018172065 A1.

As the protein of the invention is preferably produced by expression in plants or cells thereof, the composition may be a plant material or extract thereof, wherein the plant material is a material from a plant having expressed the protein, preferably *Nicotiana* or an edible plant having expressed said protein. An extract of plant material is an aqueous solution containing water-soluble proteins including a protein of the invention that is present or expressed in said plant material, or a dried product of such aqueous solution. The extract preferably has water-insoluble components of the plant material removed e.g. by filtration or centrifugation. The plant material may be a material from a plant selected from the group consisting of spinach, chard, beetroot, carrot, sugar beet, leafy beet, amaranth, *Nicotiana*, and/or said plant material is one or more leaves, roots, tubers, or seeds, or a crushed, milled or comminuted product of said leaves, roots, tubers, or seeds.

The composition or said extract from a plant material may be a solid or liquid composition, such as a solution or a dispersion, containing the klebicin(s) of the invention. The liquid composition may be aqueous, such as an aqueous solution. The concentration of said protein in said aqueous dispersion or solution may be from 0.0001 to 1 mg/ml, preferably from

0.001 to 0.1 mg/ml, more preferably from 0.005 to 0.05 mg/ml. If more than one klebicin capable of exerting a cytotoxic effect on *Klebsiella* is employed, these concentrations relate to the total concentration of all such klebicins.

The aqueous solution may, apart from the one or more protein(s) of the invention, contain a buffer. The buffer may be an inorganic or organic acid or salts thereof. An example of an inorganic acid is phosphoric acid or salts thereof. Examples of the organic acid are HEPES, acetic acid, succinic acid, tartaric acid, malic acid, benzoic acid, cinnamic acid, glycolic acid, lactic acid, citric acid, and ascorbic acid. Preferred organic acids are malic acid, lactic acid, citric acid, and ascorbic acid. The pH of the solution may generally be from 4 to 8, preferably from 5 to 8, more preferably from 6.0 to 7.5. If the object to which the composition is applied to is meat, the pH of the solution may generally be from 4 to 8, preferably from 4.5 to 7, more preferably from 5.0 to 6.5, and even more preferably from 5.0 to 6.0. Further, the solution may contain isotonic agents such as glycerol or a salt. A preferred salt to be used is sodium chloride. The aqueous solution containing the one or more klebicin(s) may be a buffered aqueous solution that may contain further solutes e.g. salts such as from 50 to 400 mM NaCl, preferably from 100 to 200 mM NaCl. The aqueous solution may further contain a sulfhydryl compound such as dithiothreitol (DTT), dithioerythritol, thioethanol or glutathione, preferably DTT. The concentration of the total of sulfhydryl compounds in the aqueous solution may be from 1 to 50 mM, preferably from 2 to 20 mM and more preferably from 4 to 10 mM.

If the composition of the invention is a solid composition, it may be a powder such as a lyophilized solid composition obtained by lyophilization of the extract or solution mentioned above. The powder may contain additional solid components such as those mentioned above for the aqueous solution. Before use, it may be reconstituted with a suitable liquid, such as water or buffer. The solid composition may contain buffer, salts or other components as mentioned above, such that the concentrations given above may be achieved upon reconstitution or dissolution of the solid composition.

Examples of carriers of the composition are solvents such as water or an aqueous buffer (as described above), salts, sugars such as monosaccharides and disaccharides, sugar alcohols, and other carriers such as those known from pharmaceutical compositions. Examples of the latter are starch, cellulose and other proteins such as albumin. Examples of sugars are glucose, fructose, lactose, sucrose, and maltose.

The composition of the invention may contain at least 10, preferably at least 20, more preferably at least 30, even more preferably at least 50, even more preferably at least 75 % by weight of one or more klebicin(s) of the invention based on the total weight of protein in the composition. The content of klebicin(s) in the composition may be determined by subjecting the composition to SDS-PAGE and analyzing the obtained gel, after staining, by determining the intensity of bands on the gel, according to Reference Example 1. Thereby, intensity of bands

due to klebicins can be determined in relation to the sum of intensities of bands due to all proteins in the composition.

In one embodiment, the composition of the invention is a pharmaceutical composition. The pharmaceutical composition may, apart from one or more proteins of the invention, optionally contain an *E. coli* colicin. It also contains one or more suitable pharmaceutically acceptable carrier and/or excipients, depending on whether it is liquid or solid and depending on the intended use. The excipients or carrier may be those mentioned above.

The composition of the invention such as the pharmaceutical composition may be formulated for oral delivery to the small or large intestine. Thus, the invention also provides an oral formulation comprising the protein of the invention or the composition according to the invention said formulation being capable of protecting the protein from gastric conditions (e.g. acidic pH and/or proteases) and capable of releasing the protein in the intestine. The capability of protecting the protein from gastric conditions and of releasing the protein in the intestine preferably relates to that in a mammal, preferably in a human subject. Enteric delivery of drugs or their active ingredients for avoiding the degradation of the drug or active ingredient by the acidic gastric conditions or gastric proteolytic conditions is known to the skilled person. Solid compositions or formulations such as tablets may be coated with a polymer that is resistant to gastric conditions but dissolves under the more neutral conditions of the intestine. Examples of commercial products suitable for coating are Eudragit™ S100 from Evonik or enTRinsic™ drug-delivery technology from Lonza Company.

In another embodiment, the composition of the invention such as the pharmaceutical composition may be formulated for delivery to the lungs. Thus, the invention also provides a pulmonary formulation comprising the protein of the invention or the composition according to the invention. For an overview over topical lung delivery of protein therapeutics see e.g. Bodier-Montagutelli et al., EXPERT OPINION ON DRUG DELIVERY 2018, VOL. 15, NO. 8, 729–736; doi.org/10.1080/17425247.2018.1503251. The formulation may be a dry powder for aerosolization or a liquid solution for nebulization.

Other possible embodiments of formulations of the composition are given below in the section on medical applications.

Application to objects

The invention provides a method of preventing or reducing contamination of an object with *Klebsiella*, comprising contacting said object with one or more proteins (klebicins) as described above or a composition as described above. The object is a non-living object. The object may be a surface of any non-organic object or an organic object such as food. Contamination of an object with *Klebsiella* means adhesion of viable *Klebsiella* cells to the object. Reducing contamination with *Klebsiella* means reducing the number of viable *Klebsiella*

cells adhering to the object. Determining contamination of objects with *Klebsiella* is part of the general knowledge. For example, dilution plating of solutions or dispersions of homogenized food as done in the Examples or dilution plating of a rinsing solution of other objects may be used, followed by counting bacterial colonies. Preferably, the object is food or animal feed.

For treating or contacting the object with the protein or composition of the invention, a solution of the protein or a liquid composition as described above is generally contacted with the object. For example, said object is sprayed with an aqueous solution or is immersed into the aqueous solution as a composition of the invention. The object may be immersed for at least 10 seconds, preferably for at least 1 minute, preferably for at least 5 minutes into the aqueous solution. Contacting the object with a liquid composition helps to distribute the composition over the surface of the object. Where sufficiently even distribution can be achieved, it is possible to contact the object with a solid composition according to the invention.

Medical applications

The invention also provides the protein, composition or pharmaceutical composition of the invention for use in the treatment or prevention of an infection of a subject with *Klebsiella*, notably the *Klebsiella* species mentioned above. The invention also provides a method of treating or preventing infection of a subject with *Klebsiella*, notably the Klebsiella species mentioned above, comprising administering to said subject one or more proteins (klebicin(s)) or the composition of the invention. The subject may be a human being or a mammal such as a farm animal. Human subjects are preferred. The infection to be treated may be an infection by *Klebsiella* that is antibiotic-resistant. The resistance may be carbapenem-resistance or a multi-drug resistance.

The *Klebsiella* infection to be treated may be any of the *Klebsiella* species described above. Klebicins KpneM (SEQ ID NO:1) and KvarM (SEQ ID NO:2) and proteins defined herein using SEQ ID NO:1 or SEQ ID NO:2 as a reference sequence are preferred proteins for the medical applications due to their wide activity against many different isolates of *Klebsiella* as demonstrated in the Examples below. Therefore, these klebicins are preferably used for treating infection (as well as for preventing or reducing contamination, see above) with any *Klebsiella*, notably of *Klebsiella pneumoniae*.

The infection by *Klebsiella* to be treated may, for example, be an infection of the urinary tract, lower respiratory tract, biliary tract, surgical wounds, or syndromes (clinical syndromes) including pneumonia, bacteremia, thrombophlebitis, cholecystitis, diarrhea, upper respiratory tract infection, osteomyelitis, and meningitis, preferably pneumonia, bacteremia, thrombophlebitis, urinary tract infection (UTI), diarrhea, upper respiratory tract infection, and wound infection.

Generally, a liquid or solid pharmaceutical composition containing the klebicin(s) and optionally further components as described above is prepared for administration to the subject. Liquid compositions may be aqueous solutions as described above. Solid compositions may be powder containing at least one klebicin, e.g. in freeze-dried form, or tablets obtained from such powder or capsules filled with such powder.

The route of administration of the protein or pharmaceutical composition depends on the disease to be treated. For the treatment of diarrhea and upper respiratory tract infections, administration may be oral, e.g. in the form of a tablet or a solution. For the treatment of diarrhea, the pharmaceutical preparation may be one that allows passage through the stomach without being attacked by the acid medium in the stomach. The klebicin(s) should then be released from the pharmaceutical composition in the intestine. Such pharmaceutical preparations are known in the art. Examples are tablets and capsules resistant to the acid medium in the stomach. It is further possible to administer orally a biological material such as *E. coli* or plant material containing expressed klebicin(s) to a patient.

For the treatment of pneumonia, e.g. if the subject has a lung infection with *Klebsiella pneumoniae*, the pharmaceutical preparation may be administered to a subject as an aerosol or powder to the lungs. Methods for formulating a protein for administration to the lungs are known in the art, see e.g. inhaled recombinant DNAseI, or Dornase : Witt DM, Anderson L. Dornase alfa: a new option in the management of cystic fibrosis. Pharmacotherapy. 1996 Jan-Feb;16(1):40-8; US20150024050A1: Dry powder formulations of Dnase I). Also see for review Depreter et al. 2013; Bodier-Montagutelli et al. 2018.

For the treatment of wound infection, the protein or pharmaceutical composition may be topically administered, e. g. as an aqueous solution. For urinary tract infection (UTI), the protein or pharmaceutical composition may be administered in the form of an aqueous solution using a catheter. For the treatment of cholecystitis or bile duct infection by *Klebsiella*, the protein or pharmaceutical composition may be administered in the form of an aqueous solution using a catheter.

The klebicin(s) may be administered to a human adult in amounts of 1 mg to 1000 mg per day, preferably of from 10 mg to 250 mg per day to a human patient. Such amounts may also be administered to an animal. In a probiotic approach, a patient may be treated by administering to the patient a genetically-modified microorganism expressing at least one of the klebicin(s). The genetically-modified microorganism may be a genetically-modified non-pathogenic *E. coli* or a lactic acid-producing microorganism as commonly employed in fermentation of milk products. Examples of lactic acid-producing microorganism are bacteria from the genera *Lactobacillus* such as *Lactobacillus lactis* and *Bifidobacterium* such as *Bifidobacterium bifidum or Bifidobacterium breve*. Another route of administration is by injection

into the blood stream of a patient for preventing infection with *Klebsiella*. For this purpose, the klebicin(s) may be dissolved in a physiological saline and the solution be sterilized.

Production of proteins of the invention

A klebicin or protein according to the invention may be produced by known methods of protein expression in a standard expression system. For producing the klebicin, a nucleotide sequence encoding it may be expressed in a suitable host organism. Methods usable for producing and purifying a protein of interest have been described in the prior art and any such methods may be used. An *E. coli* expression system as generally known in the art may, for example, be used. If a eukaryotic expression system is used, one or more introns may be inserted in the coding sequence of the klebicin to prevent toxicity on the bacterial organism used for cloning.

Particularly efficient expression methods are plant expression systems that are also known in the prior art. Plant expression systems usable for expressing a klebicin according to the invention are described in the Examples. A possible way of achieving expression of a nucleotide sequence of interest in plants is the use of self-replicating (viral) replicons containing the nucleotide sequence encoding the klebicin. The coding sequence of the klebicin may be codon optimized for expression in plants or in the particular plant used as expression host. Plant viral expression systems have been described in many publications, such as in WO2012019660, WO2008028661, WO2006003018, WO2005071090, WO2005049839, WO2006012906, WO02101006, WO2007137788 or WO02068664 and many more publications are cited in these documents. Various methods for introducing a nucleic acid molecule, such as a DNA molecule, into a plant or plant part e.g. for transient expression are known. Agrobacteria may be used for transfecting plants with the nucleic acid molecule (vector) or nucleic acid construct e.g. by agroinfiltration or spraying with agrobacterial suspensions. For references, see WO 2012019660, WO 2014187571, or WO 2013149726. The nucleic acid molecule contains a nucleotide sequence encoding a protein of the invention.

In embodiments, wherein strong expression of a klebicin as a protein of interest is desired, a nucleic acid molecule or nucleic acid construct containing a nucleotide sequence encoding the klebicin may encode a viral vector that can replicate in plant cells to form replicons of the viral vector. To replicate, the viral vector and the replicons may contain an origin of replication that can be recognized by a nucleic acid polymerase present in plant cells, such as by the viral polymerase expressed from the replicon. In case of RNA viral vectors (referred to as “RNA replicons”), the replicons may be formed by transcription under the control of a promoter active in plant cells, from the DNA construct after the latter has been introduced into plant cell nuclei. In case of DNA replicons, the replicons may be formed by recombination

between two recombination sites flanking the sequence encoding the viral replicon in the DNA construct, e.g. as described in WO00/17365 and WO 99/22003. If the replicon is encoded by the DNA construct, RNA replicons are preferred. Use of DNA and RNA viral vectors (DNA or RNA replicons) has been extensively described in the literature over the years. Some examples are the following patent publications: WO2008028661, WO2007137788, WO 2006003018, WO2005071090, WO2005049839, WO02097080, WO02088369, WO02068664. Examples of DNA viral vectors are those based on geminiviruses. For the present invention, viral vectors or replicons based on plant RNA viruses, notably those based on plus-sense single-stranded RNA viruses may be preferably used. Accordingly, the viral replicon may be a plus-sense single-stranded RNA replicon. Examples of such viral vectors are those based on tobacco mosaic virus (TMV) and potexvirus X (PVX). "Based on" means that the viral vector uses the replication system such as the replicase and/or other proteins involved in replication of these viruses. Potexvirus-based viral vectors and expression systems are described in EP2061890 or WO2008/028661. As is known from the references cited, RNA replicons such as plus-sense single-stranded RNA replicons may express a nucleotide sequence under the control of a subgenomic promoter located upstream of the nucleotide sequence. By the action of the viral replicase that may be encoded on the same RNA replicon, subgenomic RNA may be replicated in plant cells, containing the (RNA) nucleotide sequence, whereby the protein may be translated from the subgenomic RNA.

The klebicin may be expressed in a multi-cellular plant or a part thereof, notably a higher plant or parts thereof. Both monocot and dicot (crop) plants can be used. Common plants usable for expressing the protein of interest include *Nicotiana benthamiana*, *Nicotiana tabacum*, spinach, *Brassica campestris*, *B. juncea*, beets (*Beta vulgaris*), cress, arugula, mustard, strawberry, *Chenopodium capitatum*, lettuce, sunflower, cucumber, chinese cabbage, cabbage, carrot, green onion, onion, radish, lettuce, field peas, cauliflower, broccoli, burdock, turnip, tomato, eggplant, squash, watermelon, prince melon, and melon. Preferred plants are spinach, chard, beetroot, carrot, sugar beet, *Nicotiana tabacum*, and *Nicotiana benthamiana*. Expression in edible plants may be used for preventing contamination of the plants or food made therefrom with *Klebsiella*. In one embodiment, plants are used that do not normally enter the human or animal food chain such as *Nicotiana* species such as *N. tabacum* and *N. benthamiana*.

Generally, the klebicin as a protein of interest is expressed in the cytosol of cells of the plants or plant parts. In this case, no signal peptide directing the protein of interest into a particular compartment is added to the protein. Alternatively, the protein of interest can be expressed in or targeted into chloroplasts of the plants; in the latter case, an N-terminal pre-sequence, generally referred to as plastid transit peptide or chloroplast targeting peptide, is

added to the N-terminal or C-terminal end, preferably the N-terminal end, of the klebicin as the protein of interest.

The invention provides a nucleic acid molecule containing a nucleotide sequence encoding a protein of the invention. The nucleic acid molecule may comprise a nucleic acid construct that contains (i) a transcription promoter active in plant cells and (ii) a nucleotide sequence encoding a protein of the invention for expressing said nucleotide sequence in plant cells under the control of the promoter.

The invention also provides a nucleic acid molecule or nucleic acid construct encoding the protein of the invention, said nucleic acid molecule or nucleic acid construct comprising (i) a transcription promoter that is preferably active in plant cells and (ii) a nucleotide sequence encoding said protein for expressing said nucleotide sequence in cells, preferably in plant cells, under the control of said promoter.

Further, the invention provides a nucleic acid molecule or nucleic acid construct encoding a protein of the invention, said nucleic acid molecule or nucleic acid construct is or encodes a viral (DNA or RNA) replicon comprising a nucleotide sequence encoding said protein for expressing said nucleotide sequence in cells, preferably in plant cells; said replicon may contain a subgenomic promoter for expressing said nucleotide sequence in plant cells or cells of a plant under the control of said subgenomic promoter.

As the protein of the invention is preferably expressed in a plant or in plant cells, the invention also provides a plant, plant tissue, or plant cell, comprising a protein of the invention. The invention also provides a plant, plant tissue, or plant cell, comprising the nucleotide sequence or the nucleic acid molecule of the invention. The plant may be any one of those mentioned above.

Production of the composition of the invention

In the process of producing a composition comprising at least one klebicin, a klebicin is, in the first step, expressed in a plant or cells of a plant, such as an edible plant. In the next step, plant material containing expressed klebicin from a plant having expressed the klebicin is harvested. Plant material may e.g. be leaves, roots, tubers, or seeds, or a crushed, milled or comminuted product of leaves, roots, tubers, or seeds. In step (iii), the klebicin is extracted from the plant material using an aqueous buffer. This may include that the plant material is homogenized, and insoluble material may be removed by centrifugation or filtration. Soluble components including the klebicin will be extracted into the aqueous buffer to produce a klebicin solution in the aqueous buffer. The aqueous buffer may contain an inorganic or organic acid or salts thereof and may have a pH as defined above for the aqueous solution as a composition of the invention. Further, the aqueous buffer may contain salt and/or a sulfhydryl

compound as also described above for the aqueous solution as a composition of the invention. If a relatively pure klebicin composition is desired, the klebicin solution in the aqueous buffer may be further purified by removing undesired components according to known methods of protein purification.

Accordingly, the invention provides a process of producing a composition comprising a protein according to the invention, said process comprising the following steps:

(i) expressing said protein in a plant as described above, preferably an edible plant or *Nicotiana*,

(ii) harvesting plant material containing expressed protein from said plant,

(iii) extracting said protein from said plant material using an aqueous buffer to obtain a composition containing said protein,

optionally removing undesired contaminants from said composition.

If a klebicin is expressed in plants, the plants or tissue thereof having expressed protein is harvested, the tissue may be homogenized, and insoluble material may be removed by centrifugation or filtration. If relatively pure klebicin is desired, the klebicin may be further purified by generally known method of protein purification such as by chromatographic methods which can remove other host-cell proteins and plant metabolites such as alkaloids and polyphenols. Purified klebicin solutions may be concentrated and/or freeze-dried.

If klebicin are expressed in edible plants, crude protein extracts from the edible plants or semi-purified concentrates may be used for preventing or reducing contamination of an object such as food with *Klebsiella*.

## EXAMPLES

**Reference Example 1**

Soft-agar overlay assay for evaluation of klebicin toxicity

Overnight *Klebsiella quasipneumoniae* subsp. *similipneumoniae* SB30 (DSM 28212) culture is equalized to OD595=1.0 in LB medium and diluted 100x in 0.8% top agar preheated in a 55 °C water bath. Mixed overlay components are poured on plates containing solid agar (1.5% LB agar); the plates are kept for a few minutes allowing the agar to harden. Sterile Whatman discs (6 mm diameter) are placed on soft-agar and 20 µl aliquots of klebicin solution containing 10 µg klebicin protein are applied to the disks. The plates are incubated for 16 hours at 37 °C. After 16 hours incubation, the diameter of klebicin inhibition zones is measured.

Determination of klebicin concentration

The klebicin concentration in liquid sample containing the klebicin is determined by performing SDS-PAGE with Coomassie staining and reading out the intensity of the band due

to the klebicin using a commercial reader and by comparing the determined intensity with bands obtained by performing SDS-PAGE with Coomassie staining of serial dilutions of Bovine Serum Albumin (BSA) of known concentrations. A calibration curve may be obtained from the intensities of bands on stained SDS-PAGE gels of BSA. The concentration of BSA is determined using the Bradford protein assay (For example, Bradford reagent, B6916, Sigma-Aldrich, St-Louis, MO, USA).

**Example 1: Construction of klebicin expression vectors**

The KpneA (*K. pneumoniae* SAV78255.1), KaerA (*K. aerogenes* WP_063414841.1), KoxyY (*K. oxytoca* WP_024273778), Kvarla (*K. variicola* KDL88409), Kpnela (*K. pneumoniae* BAS34675), KpneM (*K. pneumoniae* EWD35590.1), KpneM2 (*Klebsiella sp.* WP_047066220), KvarM (*K. variicola* CTQ17225.1), KaerM (*K. aerogenes* WP_015367360.1) optimized for expression in the host plant *Nicotiana benthamiana* were synthetized by Thermofisher Scientific (USA) and inserted as *Bsa*I-*Bsa*I fragments in pICH29912, assembled TMV-based magnICON® vector (Marillonnet et al., 2005) (Fig. 1). Obtained plasmids were used to transform *A. tumefaciens* GV3101.

**Example 2: Expression of klebicins in plants**

*N. benthamiana* plants were grown in a growth chamber at 25 °C and 50 % humidity, with a 16 h light (1500 lux) and 8 h dark photoperiod. Four-to-six-week-old plants were used for transfection with recombinant *A. tumefaciens*.

*A. tumefaciens* were grown overnight at 30 °C in LB medium containing 50 mg $L^{-1}$ rifampicin and 50 mg $L^{-1}$ kanamycin. *Agrobacterium* overnight cultures were sedimented at 3220 *g* for 5 min and resuspended in tap water at an $OD_{595}$ of 1.5.

Four-to-six-week-old plant leaves were infiltrated into the abaxial side of the leaf using a syringe without a needle with a 1:1000 dilution of *A. tumefaciens* strain containing expression vector. Plant leaves were observed and collected at 4-7 dpi (days post infiltration).

The SDS-PAGE and Coomassie staining analysis of the extracts of soluble proteins of infiltrated plant leaves revealed that all nine klebicins are efficiently expressed in plants and are detected in the gel as very intense supplementary bands (Fig. 2). The weights of polypeptides observed in electrophoresis approximately correspond to the expected theoretical molecular weights (Kvarla – 43.4 kDa, Kpnela – 48.5 kDa, KpneA – 40 kDa, Kaer A – 39 kDa, KoxyY – 48.7 kDa, KpneM – 30.3 kDa, KpneM2 - 29.7 kDa, KvarM – 29.8 kDa and KaerM – 29 kDa). The expression level of individual klebicins varies in a range of 2.7 – 4.4 mg/g FW, the highest expression levels achieved by the two *K. pneumoniae* M-type klebicins KpneM2 and KpneM (Table 1).

**Example 3: Purification of klebicins from plant biomass**

KpneA, KaerA, Kvarla, KpneM, KpneM2 and KvarM bacteriocins were purified to homogeneity by protein chromatography. Quite pure KpneM, KpneM2 and KvarM proteins were obtained after single step hydrophobic interaction chromatography (HIC), but for best results a second purification step by anion chromatography was included. KpneA and Kvarla were also purified by using as a first step hydrophobic interaction chromatography, but then followed by cation exchange chromatography column. KaerA was purified by two steps of ion exchange chromatography, cation exchange column as a first step and anion exchange column as second step.

Crude protein extract was prepared as follows. A small portion of frozen leaf tissue was ground into fine powder with mortar and pestle using liquid nitrogen. Prepared powder was mixed with cold extraction buffer at a ratio of 1 g of plant material to 5 mL of buffer. The suspension was kept on ice for 15-20 min. Cell debris were removed by centrifugation at 3220 g at 4 °C for 20 min., and the supernatant was filtered through membrane filters (pore sizes 5 µm and 0.22 µm). Obtained solution was taken as total soluble protein and applied for purification by two-step chromatography. Details of purification protocols varied depending on proteins.

**KpneM** was purified using the combination of Hydrophobic Interaction Chromatography (HIC) and Anion Exchange Chromatography (AEXC) (Fig. 3A, B).

A small portion of frozen leaf tissue was homogenized with chilled mortar and pestle in liquid nitrogen. Prepared powder was mixed with cold extraction buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, 30 mM NaCl, pH 5.0) at a ratio of 1 g of plant material to 5 ml of buffer. The crude extract kept at 20-25 °C for 10-15 min. Cell debris were removed by centrifugation at 3220 *g*, at 4 °C for 20 min. Pellets were discarded and the supernatant was filtered by passing solution through membrane filters (pore sizes 5 µm and 0.45 µm). Ammonium sulphate was added up to 0.70 M and pH of solution adjusted to 6. Formed precipitate was removed by centrifugation at 3220 *g*, at 4 °C for 5 min. The supernatant was taken as total soluble protein and applied for purification in two steps.

At the first purification step the chromatography column was filled with Phenyl sepharose FF resin (GE Healthcare Life Sciences, Uppsala, Sweden) and pre-equilibrated with cold buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, 0.70 M $(NH_4)_2SO_4$, pH 6.0). Protein solution was loaded to column and the Phenyl sepharose bounded protein fraction was eluted by washing with elution buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, 0.28 M $(NH_4)_2SO_4$, pH 6.0). Collected protein fraction replaced to the diafiltrating concentrator (10 kDa) and centrifuged at 3220 *g* until the volume of protein solution decreased 8-10 folds. Concentrate was diluted up to a primary volume with buffer containing 50 mM $NaH_2PO_4/Na_2HPO_4$ (pH 8.0). Procedure was repeat till conductivity decreased below 10

mS/cm and protein solution subjected to the final purification step using Q sepharose FF resin (GEHealthcare Life Sciences, Uppsala, Sweden). Chromatography media was pre-equilibrated with cold buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, pH 8.0). Protein solution was loaded to column and Q sepharose unbounded protein was collected in flow through fraction. After KpneM was freeze-dried and applied for analysis.

**KpneM2** was purified using the combination of Hydrophobic Interaction Chromatography (HIC) and Anion Exchange Chromatography (AEXC) (Fig. 3C, D).

A small portion of frozen leaf tissue was homogenized with chilled mortar and pestle in liquid nitrogen. Prepared powder was mixed with cold extraction buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, 30 mM NaCl, pH 5.0) at a ratio of 1 g of plant material to 5 ml of buffer. The crude extract kept at 20-25 °C for 10-15 min. Cell debris were removed by centrifugation at 3220 *g*, at 4 °C for 20 min. Pellets were discarded and the supernatant was filtered by passing solution through membrane filters (pore sizes 5 µm and 0.45 µm). Ammonium sulphate was added up to 0.70 M and pH of solution adjusted to 6. Formed precipitate was removed by centrifugation at 3220 *g*, at 4 °C for 5 min. The supernatant was taken as total soluble protein and applied for purification in two steps.

At the first purification step the chromatography column was filled with Phenyl sepharose FF resin (GE Healthcare Life Sciences, Uppsala, Sweden) and pre-equilibrated with cold buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, 0.70 M $(NH_4)_2SO_4$, pH 6.0). Protein solution was loaded to column and the Phenyl sepharose bounded protein fraction was eluted by washing with elution buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, 0.42 M $(NH_4)_2SO_4$, pH 6.0). Collected protein fraction replaced to the diafiltrating concentrator (10 kDa) and centrifuged at 3220 *g* until the volume of protein solution decreased 8-10 folds. Concentrate was diluted up to a primary volume with buffer containing 50 mM $NaH_2PO_4/Na_2HPO_4$ (pH 8.0). Procedure was repeat till conductivity decreased below 10 mS/cm and protein solution subjected to the final purification step using Q sepharose FF resin (GEHealthcare Life Sciences, Uppsala, Sweden). Chromatography media was pre-equilibrated with cold buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, pH 8.0). Protein solution was loaded to column and Q sepharose unbounded protein was collected in flow through fraction. After KpneM2 was freeze-dried and applied for analysis.

**KvarM** was purified using the combination of Hydrophobic Interaction Chromatography (HIC) and Anion Exchange Chromatography (AEXC) (Fig. 3E, F).

A small portion of frozen leaf tissue was homogenized with chilled mortar and pestle in liquid nitrogen. Prepared powder was mixed with cold extraction buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, pH 5.0) at a ratio of 1 g of plant material to 5 ml of buffer. The crude extract kept at 20-25 °C for 10-15 min. Cell debris were removed by centrifugation at 3220 *g*, at 4 °C for 20 min. Pellets were discarded and the supernatant was filtered by passing solution through

membrane filters (pore sizes 5 µm and 0.45 µm). Ammonium sulphate was added up to 0.95 M and pH of solution adjusted to 6. Formed precipitate was removed by centrifugation at 3220 *g*, at 4 °C for 5 min. The supernatant was taken as total soluble protein and applied for purification in two steps.

At the first purification step the chromatography column was filled with Phenyl sepharose FF resin (GE Healthcare Life Sciences, Uppsala, Sweden) and pre-equilibrated with cold buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, 0.95 M $(NH_4)_2SO_4$, pH 6.0). Protein solution was loaded to column and the Phenyl sepharose bounded protein fraction was eluted by washing with elution buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, 0.62 M $(NH_4)_2SO_4$, pH 6.0). Collected protein fraction was placed in the diafiltrating concentrator (10 kDa) and centrifuged at 3220 *g* until the volume of protein solution decreased 8-10 folds. Concentrate was diluted up to a primary volume with buffer containing 50 mM $NaH_2PO_4/Na_2HPO_4$ (pH 8.0). Procedure was repeated until conductivity decreased below 10 mS/cm and protein solution was subjected to the final purification step using Q sepharose FF resin (GEHealthcare Life Sciences, Uppsala, Sweden). Chromatography media was pre-equilibrated with cold buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, pH 8.0). Protein solution was loaded to column and Q sepharose unbounded protein was collected in flow through fraction. After that, KvarM was freeze-dried and applied for analysis.

**KpneA** was purified using the combination of Hydrophobic Interaction Chromatography (HIC) and Cation Exchange Chromatography (CEXC) (Fig. 3G, H).

A small portion of frozen leaf tissue was homogenized with chilled mortar and pestle in liquid nitrogen. Prepared powder was mixed with cold extraction buffer (20 mM $NaH_2PO_4/Na_2HPO_4$, 30 mM NaCl, pH 5.0) at a ratio of 1 g of plant material to 5 ml of buffer. The crude extract kept at 20-25 °C for 10-15 min. Cell debris were removed by centrifugation at 3220 *g*, at 4 °C for 20 min. Pellets were discarded and the supernatant was filtered by passing solution through membrane filters (pore sizes 5 µm and 0.45 µm). Ammonium sulphate was added up to 1.50 M and pH of solution adjusted to 6. Formed precipitate was removed by centrifugation at 3220 *g*, at 4 °C for 5 min. The supernatant was taken as total soluble protein and applied for purification in two steps.

At the first purification step the chromatography column was filled with Phenyl sepharose FF resin (GE Healthcare Life Sciences, Uppsala, Sweden) and pre-equilibrated with cold buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, 1.50 M $(NH_4)_2SO_4$, pH 6.0). Protein solution was loaded to column and the Phenyl sepharose bounded protein fraction was eluted by washing with elution buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, 0.90 M $(NH_4)_2SO_4$, pH 6.0). Collected protein fraction was placed in the diafiltrating concentrator (10 kDa) and centrifuged at 3220 *g* until the volume of protein solution decreased 8-10 folds. Concentrate was diluted up to a primary volume with buffer containing 20 mM $NaH_2PO_4/Na_2HPO_4$, 20 mM Sodium citrate (pH 4.5). Procedure was repeated till conductivity

decreased below 9 mS/cm and protein solution was subjected to the final purification step using SP sepharose FF resin (GEHealthcare Life Sciences, Uppsala, Sweden). Chromatography media was pre-equilibrated with cold buffer (20 mM $NaH_2PO_4/Na_2HPO_4$, 20 mM Citric acid, pH 4.5). Protein solution was loaded to column and SP sepharose bounded protein fraction was eluted by linear gradient of cold washing buffer additionally containing 500 mM of NaCl. After that, KpneA was freeze-dried and applied for analysis.

**KaerA** was purified using the combination of Cation Exchange Chromatography (CEXC) and Anion Exchange Chromatography (AEXC) (Fig. 3I, J).

A small portion of frozen leaf tissue was homogenized with chilled mortar and pestle in liquid nitrogen. Prepared powder was mixed with cold extraction buffer (20 mM $NaH_2PO_4/Na_2HPO_4$, 20 mM Citric acid, pH 5.0) at a ratio of 1 g of plant material to 5 ml of buffer. The crude extract kept at 20-25 °C for 10-15 min. Cell debris were removed by centrifugation at 3220 *g*, at 4 °C for 20 min. Pellets were discarded and the supernatant was filtered by passing solution through membrane filters (pore sizes 5 µm and 0.45 µm). The pH of solution was adjusted to 4.5 and formed precipitate was removed by centrifugation at 3220 *g*, at 4 °C for 5 min. The supernatant was taken as total soluble protein and applied for purification in two steps.

At the first purification step the chromatography column was filled with SP sepharose FF resin (GE Healthcare Life Sciences, Uppsala, Sweden) and pre-equilibrated with cold buffer (20 mM $NaH_2PO_4/Na_2HPO_4$, 20 mM Citric acid, pH 4.5). Protein solution was loaded to column and SP sepharose bounded protein fraction was eluted by linear gradient of cold washing buffer additionally containing 500 mM of NaCl. Collected protein fraction was placed in the diafiltrating concentrator (10 kDa) and centrifuged at 3220 *g* until the volume of protein solution decreased 8-10 folds. Concentrate was diluted up to a primary volume with buffer containing 20 mM $NaH_2PO_4/Na_2HPO_4$, (pH 8.0). Procedure was repeated until conductivity decreased below 8 mS/cm and protein solution was subjected to the final purification step using Q sepharose FF resin (GEHealthcare Life Sciences, Uppsala, Sweden). Chromatography media was pre-equilibrated with cold buffer (20 mM $NaH_2PO_4/Na_2HPO_4$, pH 8.0). Protein solution was loaded to column and Q sepharose unbounded protein was collected in flow through fraction. After that, KaerA was freeze-dried and applied for analysis.

**Kvarla** was purified using the combination of Hydrophobic Interaction Chromatography (HIC) and Cation Exchange Chromatography (CEXC) (Fig. 3K, L).

A small portion of frozen leaf tissue was homogenized with chilled mortar and pestle in liquid nitrogen. Prepared powder was mixed with cold extraction buffer (20 mM $NaH_2PO_4/Na_2HPO_4$, 30 mM NaCl, pH 5.0) at a ratio of 1 g of plant material to 5 ml of buffer. The crude extract kept at 20-25 °C for 10-15 min. Cell debris were removed by centrifugation at 3220 *g*, at 4 °C for 20 min. Pellets were discarded and the supernatant was filtered by passing solution

through membrane filters (pore sizes 5 µm and 0.45 µm). Ammonium sulphate was added up to 1.35 M and pH of solution adjusted to 6. Formed precipitate was removed by centrifugation at 3220 *g*, at 4 °C for 5 min. The supernatant was taken as total soluble protein and applied for purification in two steps.

At the first purification step the chromatography column was filled with Phenyl sepharose FF resin (GE Healthcare Life Sciences, Uppsala, Sweden) and pre-equilibrated with cold buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, 1.35 M $(NH_4)_2SO_4$, pH 6.0). Protein solution was loaded to column and the Phenyl sepharose bounded protein fraction was eluted by washing with elution buffer (50 mM $NaH_2PO_4/Na_2HPO_4$, 0.81 M $(NH_4)_2SO_4$, pH 6.0). Collected protein fraction was placed in the diafiltrating concentrator (10 kDa) and centrifuged at 3220 *g* until the volume of protein solution decreased 8-10 folds. Concentrate was diluted up to a primary volume with buffer containing 20 mM $NaH_2PO_4/Na_2HPO_4$, 20 mM Sodium citrate (pH 4.5). Procedure was repeated till conductivity decreased below 8 mS/cm and protein solution subjected to the final purification step using SP sepharose FF resin (GEHealthcare Life Sciences, Uppsala, Sweden). Chromatography media was pre-equilibrated with cold buffer (20 mM $NaH_2PO_4/Na_2HPO_4$, 20 mM Citric acid, pH 4.5). Protein solution was loaded to column and SP sepharose bounded protein fraction was eluted by linear gradient of cold washing buffer additionally containing 500 mM of NaCl. After that, Kvarla was freeze-dried and applied for analysis.

Concentration of purified proteins was evaluated by Bradford assay or by comparison of band intensity with known BSA amount which was run on the same SDS-PAGE gel. The results of klebicin purification are summarized in the Table 1. Figure 4 shows purified KpneM, KpneM2, KvarM, KpneA, KaerA and Kvarla klebicin proteins loaded on the same gel. All purified klebicins contain only 0.2 – 3.7% of impurities, as determined by capillary gel electrophoresis. The yields of individual klebicins after purification are in range of 0.34 – 1.1 mg/g FW. The purification of klebicins with greatest expression levels give biggest final yields and also best quality of purified proteins.

**Table 1** Purification method, obtained yields and purity of plant-expressed klebicins.

| Klebicin | Purification method | Klebicin amount in crude extract (µg/g FW) | Yield of purified protein (µg/g FW) | Purity % |
|---|---|---|---|---|
| KpneM | Phenyl>DS>Q | 3239 ± 224 | 1134 ± 54 | 99.8 ± 0.3 |
| KpneM2 | Phenyl>DS>Q | 4448 ± 347 | 920 ± 55 | 99.5 ± 0.5 |
| KvarM | Phenyl>DS>Q | 2423 ± 146 | 535 ± 35 | 98.1 ± 1.0 |
| KpneA | Phenyl>DS>SP | 2677 ± 163 | 337 ± 26 | 97.2 ± 1.2 |
| KaerA | SP>DS>Q | 2718 ± 227 | 468 ± 46 | 96.3 ± 0.7 |
| Kvarla | Phenyl>DS>SP | 2697 ± 149 | 629 ± 31 | 98.9 ± 1.0 |

**Example 4: Klebicin activity tests in soft-agar overlay assay**

We tested the activity of crude bacteriocin-expressing plant extracts in a soft-agar overlay assay with twelve *Klebsiella* strains belonging to different species (*K. pneumoniae*, *K. quasipneumoniae*, *K. oxytoca*, *K. variicola* and *K. aerogenes*). *Klebsiella* strains were purchased from Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures and are described in Table 2.

Overnight *Klebsiella* cultures were equalized to $OD_{595}$=1.0 in LB medium and diluted 100x in 0.8% top agar preheated in a 55 °C water bath. Mixed overlay components were poured on plates containing solid agar (1.5% LB agar); the plates were kept at room temperature for a few minutes allowing the agar to harden. Sterile Whatman discs (6 mm diameter) were placed on soft-agar and respective amounts of klebicins (20 µl of crude extracts or 10 µg of purified klebicins) were applied to the disks. The plates were incubated overnight at 37 °C and the diameter of klebicin inhibition zones was observed. The results of this assay are summarized in Fig. 5.

Two of tested bacteriocins, KoxyY and KaerM demonstrated perceptible but narrow inhibition zones on the lawn of several tested strains, with KaerM forming larger hazy inhibition zones only on *K. aerogenes* lawn. Because of the weaker activity, these two proteins were not included in further experiments.

**Table 2** *Klebsiella* strains used in the study.

| Strain | Culture collection number | Growing temperature |
|---|---|---|
| *Klebsiella pneumoniae subsp.pneumoniae* | DSM 26371, ATCC 700603 | 28 °C |
| *Klebsiella pneumoniae* | DSM 789, ATCC 4352 | 37 °C |
| *Klebsiella pneumoniae subsp.pneumoniae* | DSM 9377, ATCC 13887 | 37 °C |
| *Klebsiella pneumoniae subsp. rhinoscleromatis* | DSM 16231, ATCC 13884 | 37 °C |
| *Klebsiella pneumoniae subsp.ozaenae* | DSM 16358, ATCC 11296 | 28 °C |
| *Klebsiella quasipneumoniae subsp.quasipneumoniae* | DSM 28211 | 37 °C |
| *Klebsiella quasipneumoniae subsp. similipneumoniae* | DSM 28212 | 37 °C |
| *Klebsiella oxytoca* | DSM 5175, ATCC 13182 | 37 °C |
| *Klebsiella oxytoca* | DSM 6673, ATCC 43863 | 37 °C |
| *Klebsiella variicola* | DSM 15968, ATCC BAA-830 | 28 °C |
| *Klebsiella aerogenes* | DSM 30053 | 30°C |
| *Klebsiella aerogenes* | DSM 12058 | 30°C |

All seven remaining bacteriocins formed large inhibition zones on the lawn of several tested *Klebsiella* species and strains. All twelve tested strains were inhibited by not only one, but by several bacteriocins. The three remaining colM-like proteins demonstrated the largest activity spectrum and a similar activity pattern, targeting eleven out of twelve tested strains. However,

KvarM formed significantly larger inhibition zones than both *K. pneumoniae* colM-like bacteriocins (KpneM and KpneM2). The two ColA-like proteins KpneA and KaerA also demonstrated very similar activity pattern, although zone diameter was different for some of the tested strains. And finally, both Colla-like proteins Kvarla and Kpnela demonstrated very similar activity patterns (Fig. 5). All bacteriocins formed inhibition zones on the strains belonging to all five different *Klebsiella* species with exception of Kvarla and Kpnela. The two colla-like proteins had little effect on neither of four tested *K. pneumoniae* strains.

**Example 5: Evaluation of klebicin activity against a panel of clinical *Klebsiella* isolates**

All six purified klebicins were next tested against a larger panel of Klebsiella strains: 89 *K. pneumoniae* and 11 *K. oxytoca* strains, in total one hundred clinical *Klebsiella* isolates. Clinical Klebsiella strains used for agar overlay assay have been isolated in Lithuanian university of health sciences, Kaunas clinics, and are described in Table 3. Purified lyophilized klebicins were resuspended in deionized water and applied as 10 µl drops (10 µg of protein) on 6 mm Whatman discs placed on LB plates with streaked *Klebsiella*. After overnight incubation inhibition zones were measured.

KvarM demonstrated a surprisingly broad spectrum of activity. 85 % of strains were sensitive to this klebicin (Fig. 6, Table 3). KpneM was not far behind, targeting 74 % of tested strains, in general with slightly smaller inhibition zones. The specificity of KvarM and KpneM activity spectra were largely overlapping, but KvarM targeted 11 strains more that KpneM, and only one strain immune to KvarM was sensitive to KpneM (Fig. 6, Table 3). In contrast, the third M-type klebicin KpneM2 was much less active, targeting only 20 % of strains. Both colA-like klebicins KpneA and KaerA targeted 30 % and 28 % of strains, respectively, with partially overlapping profiles. 9 strains immune to KaerA were sensitive to KpneA, and 7 strains immune to KpneA were sensitive to KaerA. KpneA also in general formed larger inhibition zones. Kvarla had the narrowest spectrum of activity and targeted only 10% of all strains, 6 *K. oxytoca* and 4 *K. pneumoniae* (Fig. 6, Table 3).

**Table 3** Antimicrobial activity of plant-made klebicins against clinical *Klebsiella* isolates. "MDR" means multi-drug-resistant.

| No inhibition zone | Zone 7-10mm | Zone 11-15mm | Zone 16-20mm |
|---|---|---|---|
| - | + | + | + |

| No | Species | Isolated from | MDR | KpneA | KaerA | KpneM | KpneM2 | KvarM | Kvarla |
|---|---|---|---|---|---|---|---|---|---|
| 1 | *K. pneumoniae* | urine | MDR | - | - | + | + | + | - |
| 2 | *K. pneumoniae* | bronch | MDR | - | - | + | + | + | - |
| 3 | *K. pneumoniae* | urine | MDR | - | - | + | + | + | - |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 4 | *K. pneumoniae* | trachea | - | - | - | + | - | + | - |
| 5 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 6 | *K. pneumoniae* | blood | MDR | - | - | + | + | + | - |
| 7 | *K. pneumoniae* | urine | - | - | - | + | - | - | - |
| 8 | *K. pneumoniae* | urine | MDR | + | + | + | + | + | - |
| 9 | *K. pneumoniae* | urine | MDR | - | - | - | - | + | - |
| 10 | *K. pneumoniae* | gall | MDR | - | + | + | - | + | - |
| 11 | *K. pneumoniae* | pleura | MDR | - | - | + | - | + | - |
| 12 | *K. pneumoniae* | bronch | - | + | - | + | - | + | - |
| 13 | *K. pneumoniae* | urine | MDR | - | - | + | + | + | - |
| 14 | *K. pneumoniae* | urine | MDR | - | + | + | - | + | - |
| 15 | *K. pneumoniae* | urine | MDR | + | - | + | + | + | - |
| 16 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 17 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 18 | *K. pneumoniae* | urine | MDR | - | - | + | + | + | - |
| 19 | *K. pneumoniae* | urine | - | + | + | + | + | + | - |
| 20 | *K. pneumoniae* | urine | MDR | + | + | + | - | + | - |
| 21 | *K. pneumoniae* | urine | MDR | + | - | + | + | + | - |
| 22 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 23 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 24 | *K. pneumoniae* | urine | MDR | - | + | + | + | + | - |
| 25 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 26 | *K. pneumoniae* | urine | MDR | + | - | + | + | + | - |
| 27 | *K. pneumoniae* | blood | MDR | - | - | + | - | + | - |
| 28 | *K. pneumoniae* | urine | MDR | - | - | - | - | + | - |
| 29 | *K. pneumoniae* | urine | MDR | - | - | + | + | + | - |
| 30 | *K. pneumoniae* | blood | MDR | + | + | + | + | + | - |
| 31 | *K. pneumoniae* | bronch | MDR | + | - | + | - | + | - |
| 32 | *K. pneumoniae* | bronch | MDR | - | - | + | - | + | - |
| 33 | *K. pneumoniae* | bronch | - | + | + | - | - | - | + |
| 34 | *K. pneumoniae* | bronch | MDR | - | - | + | - | + | - |
| 35 | *K. pneumoniae* | wound | MDR | - | - | + | - | + | - |
| 36 | *K. pneumoniae* | urine | MDR | - | - | - | - | + | - |
| 37 | *K. pneumoniae* | bronch | - | + | + | + | - | + | - |
| 38 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 39 | *K. pneumoniae* | pus | MDR | - | - | + | - | + | - |
| 40 | *K. pneumoniae* | urine | MDR | - | - | - | - | - | - |
| 41 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 41 | *K. pneumoniae* | urine | - | - | - | + | - | + | - |
| 43 | *K. pneumoniae* | bronch | MDR | - | - | + | - | + | - |
| 44 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 45 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 46 | *K. pneumoniae* | urine | MDR | - | - | - | - | + | - |
| 47 | *K. pneumoniae* | bronch | MDR | + | - | + | + | + | - |
| 48 | *K. pneumoniae* | bronch | MDR | - | - | + | - | + | - |
| 49 | *K. pneumoniae* | wound | MDR | - | - | + | + | + | - |
| 50 | *K. pneumoniae* | pus | MDR | - | - | - | - | + | + |
| 51 | *K. pneumoniae* | gall | MDR | - | - | + | - | + | - |
| 52 | *K. pneumoniae* | gall | - | - | - | + | - | + | + |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 53 | *K. pneumoniae* | gall | MDR | - | - | + | + | + | - |
| 54 | *K. pneumoniae* | urine | - | + | + | + | - | + | - |
| 55 | *K. pneumoniae* | pus | MDR | - | - | + | - | + | - |
| 56 | *K. pneumoniae* | urine | MDR | - | - | - | - | - | - |
| 57 | *K. pneumoniae* | joint | - | - | - | + | - | + | - |
| 58 | *K. pneumoniae* | phlegm | MDR | - | - | + | - | + | - |
| 59 | *K. pneumoniae* | bronch | - | + | + | + | - | + | - |
| 60 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 61 | *K. pneumoniae* | pus | - | + | + | + | + | + | - |
| 62 | *K. pneumoniae* | abdo-men | MDR | - | - | + | - | + | - |
| 63 | *K. pneumoniae* | urine | MDR | - | - | - | - | - | - |
| 64 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 65 | *K. pneumoniae* | urine | - | - | - | - | - | - | - |
| 66 | *K. pneumoniae* | blood | MDR | - | - | + | - | + | - |
| 67 | *K. pneumoniae* | pleura | MDR | - | - | + | - | + | - |
| 68 | *K. pneumoniae* | bronch | - | + | + | + | - | + | - |
| 69 | *K. pneumoniae* | urine | MDR | + | + | + | + | + | - |
| 70 | *K. pneumoniae* | blood | - | + | + | - | - | + | - |
| 71 | *K. pneumoniae* | bronch | MDR | - | - | + | - | + | - |
| 72 | *K. pneumoniae* | bronch | MDR | - | - | + | - | + | - |
| 73 | *K. pneumoniae* | urine | - | + | + | + | - | + | - |
| 74 | *K. pneumoniae* | trachea | - | + | + | + | - | + | + |
| 75 | *K. pneumoniae* | gall | - | - | - | + | - | + | - |
| 76 | *K. pneumoniae* | urine | MDR | + | + | + | - | + | - |
| 77 | *K. pneumoniae* | urine | MDR | + | - | + | - | + | - |
| 78 | *K. pneumoniae* | urine | - | + | + | + | - | + | - |
| 79 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 80 | *K. pneumoniae* | urine | MDR | - | - | - | - | - | - |
| 81 | *K. pneumoniae* | bronch | MDR | - | - | + | - | + | - |
| 82 | *K. pneumoniae* | bronch | - | - | - | - | - | - | - |
| 83 | *K. pneumoniae* | urine | - | + | + | - | - | - | - |
| 84 | *K. pneumoniae* | urine | MDR | - | - | - | - | - | - |
| 85 | *K. pneumoniae* | blood | - | + | + | + | - | + | - |
| 86 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 87 | *K. pneumoniae* | urine | MDR | - | - | - | - | - | - |
| 88 | *K. pneumoniae* | urine | MDR | - | - | + | - | + | - |
| 89 | *K. pneumoniae* | bronch | - | + | + | + | - | + | - |
| 90 | *K. oxytoca* | urine | - | + | + | - | - | + | - |
| 91 | *K. oxytoca* | urine | MDR | - | + | - | - | - | + |
| 92 | *K. oxytoca* | urine | MDR | - | + | - | - | - | + |
| 93 | *K. oxytoca* | bronch | - | - | + | + | - | + | - |
| 94 | *K. oxytoca* | urine | - | - | - | - | - | + | - |
| 95 | *K. oxytoca* | biopsy | - | - | + | - | - | + | - |
| 96 | *K. oxytoca* | urine | - | - | - | - | - | - | + |
| 97 | *K. oxytoca* | urine | - | + | + | - | - | + | + |
| 98 | *K. oxytoca* | pus | - | + | - | - | - | + | - |
| 99 | *K. oxytoca* | urine | - | + | - | - | + | + | + |
| 100 | *K. oxytoca* | bronch | - | - | - | - | - | - | + |

**Example 6: Evaluation of klebicin activity against *Klebsiella* strains in liquid cultures and biofilms**

We next performed a more detailed analysis of klebicin activity in liquid medium and on young, one day old biofilms with five representatives of different *Klebsiella* species: *K. pneumoniae*, *K. quasipneumoniae*, *K. oxytoca*, *K. variicola* and *K. aerogenes*.

For evaluation of klebicin activities in liquid medium, overnight *Klebsiella* cultures were diluted to $OD_{595}$=0.3 in iron-deficient Casamino Acids (CAA) medium (BD Bioscience) up to 1.2 mL. Lyophilized purified klebicins were resuspended in CAA medium, added to diluted bacterial suspension and incubated for 5.5-6.5 hours at 37 °C with shaking (200 rpm). The antimicrobial activity of klebicins was evaluated by determining cell numbers of bacterial test culture. Serial dilutions of 10, $10^{-1}$, $10^{-2}$, $10^{-3}$, $10^{-4}$ and $10^{-5}$ were made, plated on LB agar plates, incubated overnight at 37 °C and the CFU calculated.

Biofilms were grown as described by Moskowitz et al. (2004) and Paškevičius et al. (2017) with some modifications. Briefly, *K. quasipneumoniae, K. oxytoca, K. variicola, K. aerogenes* strains were grown overnight in LB and diluted to OD=0.1 with fresh CAA medium. 10 µl of bacteria culture were transferred to the wells of a 96-well microtiter plate (Nalgene Nunc International, Rochester, N.Y.) with 90 µl of CAA medium. Bacterial biofilms were formed by immersing the pegs of a modified polystyrene microtiter lid (Nunc TSP system) into the biofilm growth plate, followed by incubation at 30 °C or 37 °C controlled by a thermostat for 20 h. For the treatment with klebicins, peg lids were rinsed three times in sterile water, placed into microtiter plates containing 5 µg/mL of klebicins diluted in 100 µl CAA per well and incubated for 5 h at 30°C or 37 °C, depending on the strain. After incubation with klebicins, peg lids were again rinsed three times in sterile water, placed into CAA in a sterile microtiter plate and centrifuged at 810 *g* for 30 min. 6 identically treated wells were pooled each time, serial dilutions were made and bacteria were plated on LB plates for CFU counting.

For the liquid culture assay, one best performing klebicin of each group (colM-like, colA-like and colla-like) from Example 4 at concentration of 5 µg $mL^{-1}$ was tested. Klebicin KvarM inhibited the growth of all five strains, reducing the CFU number in quite similar extent - by four orders of magnitude for *K. pneumoniae* DSM 16231 and about three orders of magnitude for all remaining *Klebsiellae* (Fig. 7). Kvarla inhibited four strains and was the most efficient of all three klebicins, reducing CFU number by four to nine orders of magnitude, depending on the strain (Fig. 7) (as it was shown in Fig. 5, *K. pneumoniae* DSM16231 is insensitive to this klebicin). KpneA reduced CFU counts of three strains by 4.6 to 5.7 logs (Fig. 7).

For the biofilm assays, we used the same *Klebsiella* strains as in the liquid culture assays. First, we tested the ability of these five strains to form biofilms. Four of the tested strains formed biofilms in tested conditions, with exception of *K. pneumoniae* DSM 16231, thus this strain was not used for further experiments. The biofilms of all remaining four stains were treated for 20 h each with two klebicins, which showed the best results in liquid culture assays. The results obtained with biofilms quite closely reflected the results obtained in liquid culture assays, with exception of K. quasipneumoniae, whose biofilms were completely eradicated by KpneA and Kvarla (Fig. 8). For all remaining strains, klebicins treatment decreased CFU numbers in biofilms in similar extent as in liquid cultures, with only slight variations of Δlogs achieved (Fig. 7, Fig. 8).

**Example 7: Evaluation of klebicin antimicrobial activity *in vivo***

For the initial demonstration of klebicins activity *in vivo*, we performed *Klebsiella* challenge assay in a non-mammal animal model, *Galleria mellonella* larvae. *Galleria mellonella*, the greater wax moth or honeycomb moth, is a moth of the family *Pyralidae*. *Galleria mellonella* have been shown to be a convenient model organism for *in vivo* toxicology and pathogenicity testing, replacing the use of small mammals in such experiments (Harding et al. 2013; Paškevičius et al. 2017).

We selected Kvarla for this assay as one of the most active klebicins and K. quasipneumoniae DSM 28212 as Kvarla-sensitive challenge strain. *G. mellonella* challenge experiments were performed as described in Paškevičius et al. (2017), with some modifications. Overnight *K. quasipneumoniae* DSM 28212 strain culture was grown in CAA medium and diluted in 0.8 % NaCl in order to achieve a concentration of 1.2-3.2 x $10^6$ CFU $mL^{-1}$ in 10 µL of *K. quasipneumoniae* culture and 10 µL of klebicins solution were injected into hemocoel of fifth instar *G. mellonella* larvae (Livefood UK) in proximity of the left and/or right prolegs. Klebicins were injected two hours post infection with *K. quasipneumoniae*. Injected larvae were incubated at 37 °C in 9 cm Petri dishes without food for up to 3 days. Caterpillars were considered dead when they displayed no movement in response to mechanical stimulus to the head, leading to distinct change in color from cream to dark brown/black. Twenty larvae were used per each treatment point.

First, we determined that minimal lethal dose (MLD) of challenge strain sufficient to kill all the larvae in 68h (the duration of experiment) is 2.3 x $10^4$ CFU.

We next performed the challenge experiment with MLD and with two additional challenge doses, one inferior and one superior to MLD by factor 1.9 and 1.4, respectively. 1.2 x 104 CFU were not sufficient to kill all the larvae, as 15 % of larvae still survived after 68h. However, 2.3 and 3.2 x $10^4$ CFU were sufficient to kill all the larvae in 44 h. Injection of Kvarla 2 hours post infection completely rescued all the larvae infected by 1.2 and 2.3 x $10^4$ CFU. Larvae infected by

the highest amount of bacteria (3.2 x $10^4$ CFU) were rescued partially, with 85 % of larvae surviving till the end of experiment (Fig. 9).

Summarizing, from the activity assays, one can see that KvarM has exceptionally wide spectrum of activity as it could target *Klebsiella* strains belonging to *K. pneumoniae, K. quasipneumoniae, K. variicola, K. oxytoca,* and *K. aerogenes* species. It was also active against 85 % of strains in the panel of antibiotic-resistant clinical *Klebsiella* isolates. In liquid culture assay, this klebicin could reduce the colony forming number count by as much as three to four logs and more than by two logs in biofilm assay. Pore-forming klebicins were in general even more efficient in reducing bacterial numbers in liquid culture or in biofilms than peptidoglycan synthesis inhibitors and were able to achieve four to nine logs of CFU counts reduction in liquid cultures and two to almost six logs of CFU counts reduction in biofilms. Kvarla, which has demonstrated highest efficiency *in vitro*, was also tested *in vivo* in *G. mellonella* larvae challenge assays with very good outcome. However, the applicability of this klebicin is currently handicapped by the fact that it is not broadly active against *K. pneumoniae*, although it works well on its close relative *K. quasipneumoniae*.

**Example 8: Identification of klebicin receptors/translocators**

Universal feature of colicins is their domain organization, and each colicin appears to have receptor binding, translocation and cytotoxic domains, a feature that is conditioned by the necessity of these bacteriocins to cross the outer membrane of gram-negative bacteria (Kleanthous, 2010). Pore-forming klebicins amino acid sequence alignments with their *E. coli* counterparts reveal that their killing domains show significant degree of homology. However, as a rule, pore forming klebicins are smaller than colicins. Their amino-terminal parts, which should contain translocation and receptor binding domains, are much shorter than respective domains of colicins and have little or no sequence similarity. Thus, we anticipated that the translocation mechanism of pore-forming klebicins might be different from their *E. coli* counterparts.

In sharp contrast to some other bacteriocins, which are strictly species-specific (for example, pyocins), klebicin activity is not confined to the single species from which they are isolated, but at least to the genus. In this regard, it was important to inquire the players involved in the reception and translocation mechanisms of klebicins. *K. quasipneumoniae* DSM 28212, a strain with known genome sequence and sensitivity to all the klebicins tested, was subjected to several rounds of transposon mutagenesis and pooled mutants were tested for their sensitivity to different klebicins.

Transposon mutagenesis of *K. quasipneumoniae* DSM 28212 was performed as described in Martínez-García et al. (2011). The suicide delivery of mini-transposons localized in

pBAM1 plasmid was performed by triparental mating. The plasmid was mobilized from *E. coli* CC118λpir (pBAM1) donor cells into *K. quasipneumoniae* DSM 28212 cells with the assistance of the helper strain *E. coli* HB101 (pRK600). Obtained kanamycin-resistant clones were confirmed for the loss of the ampicillin resistance and their genomic DNA was used for the PCR amplification of the transposon adjacent regions, followed by sequencing, as described by Martínez-García et al. (2011).

29 independent mutant clones were isolated, and transposon insertions were successfully mapped in 18 klebicin-resistant mutants. To confirm that klebicin sensitivity loss was indeed due to the mapped mutations, we performed complementation assays by ectopic expression of respective wild-type genes.

For complementation assays, *Klebsiella* genome regions, containing ExbB, ExbBD, OmpC, FhuA, TonB and FimB gene ORFs along with 5' non-coding promoter regions, were PCR-amplified from *K. quasipneumoniae* DSM 28212 genomic DNA with help of Phusion DNA polymerase (Thermofisher Scientific Baltics) and ligated in pJET1.2 (Thermofisher Scientific Baltics). After sequencing, cloned fragments were excised with restriction endonucleases pair specific for each fragment, ligated in pACYC184 (NEB) and transformed into respective *K. quasipneumoniae* mutants. The sequences of primers used and cloning strategy are described in Table 4.

**Table 4** Primers used for amplification of genes used in complementation assays. Restriction endonuclease sites are in italics, primers binding sequences are in bold.

| Gene | Primer | Sequence | pACYC184 cloning |
|---|---|---|---|
| ExbB | ExbB Eco88I fwd | AAA*CTCGGG***TTGATGAACCTGTTTTTATACGTCT** (SEQ ID NO: 10) | Eco88I-Eco81I |
| | ExbB Eco81I rev | AAA*CCTGAGG***TCAACCTACCCGTAATTTCTGCG** (SEQ ID NO: 11) | |
| ExbBD | ExbB Eco88I fwd | AAA*CTCGGG***TTGATGAACCTGTTTTTATACGTCT** (SEQ ID NO: 12) | Eco88I-Eco81I |
| | ExbD Eco81I rev | AAA*CCTGAGG***TTATTTGGCTTTGACGGTCTC** (SEQ ID NO: 13) | |
| FhuA | FhuA Eco81I fwd | AAA*CCTCAGG***TTTAAGCCCTAAGACCAGACCC** (SEQ ID NO: 14) | Eco81I |
| | FhuA Eco 81I rev | AAA*CCTGAGG***TTAGAAACGGAAGGTGGCGGTG** (SEQ ID NO: 15) | |
| FimB | FimB Eco88I fwd | AAA*CTCGGG***GCTCCCGTAGCAAATAAAAACG** (SEQ ID NO: 16) | Eco88I-Eco81I |
| | FimB Eco81I rev | AAA*CCTGAGG***TTACTGAAGCAGCGACAGGCG** (SEQ ID NO: 17) | |
| OmpC | OmpC Eco88I fwd | AAA*CTCGGG***CTTGTGGCTGAACGACTCATCA** (SEQ ID NO: 18) | Eco88I-Eco81I |

| | | | |
|---|---|---|---|
| | OmpC Eco81I rev | AAA*CCTGAGG***TTAGAACTGGTAAACCAGGCCC** (SEQ ID NO: 19) | |
| TonB | TonB PsyI fwd | AAA*GACCGGGTC***GGCAAAGCTCCTTATCAATAAACA** (SEQ ID NO: 20) | BseSI-PsyI |
| | TonB BseSI rev | AAA*GTGCCC***TCAGTTAATCTCGACGCCGTTG** (SEQ ID NO: 21) | |

The summarized results from mutant klebicin sensitivity studies and complementation assays are presented in Table 5.

**Table 5** Characterization of klebicin resistant mutants obtained by transposon mutagenesis.

| Mutant No. | Selected by resistance to : | Resistant to klebicins: | Mutation | Complementation |
|---|---|---|---|---|
| **#1** | KpneM | KpneM2, KvarM, | FhuA | + |
| **#2** | KpneM | KpneM2, KvarM | FhuA | NT |
| **#3** | KpneM | KpneM2, KvarM | FhuA | NT |
| **#4** | KpneM | KpneM2, KvarM, KpneA, KaerA | TonB | + |
| **#7** | KpneM2 | KvarM, KpneM | FhuA | NT |
| **#8** | KaerA | KpneM2 (partial res. to KvarM, KpneA, KpneM) | ExbB | + (ExbBD) |
| **#9** | KaerA | Kvarla, KpneA | OmpC | + |
| **#10** | KpneA | KpneM2, KpneM, KvarM, KaerA | ExbB | + (ExbBD) |
| **#11** | KpneA | Kvarla, KaerA | OmpC | NT |
| **#12** | KvarM | KpneM, KpneM2 | FhuA | NT |
| **#13** | KvarM | KpneM, KpneM2, KaerA (partial res. to KpneA) | ExbB | NT |
| **#14** | KpneM2 | KvarM, KpneM | FhuA | NT |
| **#15** | KpneM2 | KvarM, KpneM | FhuA | NT |
| **#16** | KpneM2 | KvarM, KpneM | FhuA | NT |
| **#17** | KpneM2 | KvarM, KpneM | FhuA (outside) | NT |
| **#18** | KpneM2 | KpneM, KvarM, KaerA (partial res. to KpneA) | ExbB | NT |
| **#20** | Kvarla | KpneA, KaerA | FimB | - |

"NT" means "Not Tested", "+" means "Complemented", "-" means "Not Complemented"

Based on the obtained results, all klebicins with exception of Kvarla are similar to group B colicins and use the TonB-dependent translocation pathway. All three M-type klebicins require FhuA, TonB and ExbB for their reception–translocation, as their *E. coli* homologue colicin M. KpneA and KaerA also depend on the TonB translocation pathway and they need in addition the functional OmpC. So far we could not identify any other putative receptor for these two klebicins (Table 6).

**Table 6** Identified *Klebsiella* proteins involved in reception and translocation of klebicins.

| **Klebicin** | **Receptor** | **Mechanism of translocation** | **Cytotoxicity** |
|---|---|---|---|
| KpneM | FhuA | TonB, ExbB | Peptidoglycan synthesis inhibitor |
| KpneM2 | FhuA | TonB, ExbB | Peptidoglycan synthesis inhibitor |
| KvarM | FhuA | TonB, ExbB | Peptidoglycan synthesis inhibitor |
| KpneA | OmpC | TonB, ExbB | Pore forming |
| KaerA | OmpC | TonB, ExbB | Pore forming |
| Kvarla | OmpC | Questionable | Pore forming |

Kvarla-resistant transpositional mutants were very hard to obtain, and only some false-positive clones were isolated. Thus, we could identify only one protein participating in Kvarla reception – translocation, the outer membrane protein C (OmpC). OmpC mutants were selected by their resistance to KpneA and KaerA and it appeared that they are equally resistant to Kvarla.

Summarizing, we thus demonstrated that all three M-type klebicins KpneA, KpneM2 and KvarM are translocated by a mechanism similar to that of colicin M, and they need FhuA receptor and TonB–related translocation pathway to enter the periplasm and exercise their activity.

Two klebicins which we named KpneA and KaerA based on their killing domain similarity to colicin A, appeared also dependent on the TonB translocation pathway. This is in contrast to colicin A, which is translocated by TolA-dependent pathway. Also, while colicin A binds to BtuB, we did not isolate any BtuB mutants resistant to KpneA or KaerA. However, both KpneA and KaerA need functional OmpC, an analog of OmpF, which participates also in colA translocation (Kleanthous 2010). We have so far not identified any other putative receptor for these two klebicins.

Kvarla is different from all remaining klebicins, as it appears to be functional in all TonB and ExbB mutants. Thus, based on our results, Kvarla do not use TonB-dependent translocation pathway. Taking into account that all described colicins use either TonB or TolA as translocators, it would be expected that this protein is translocated by a Tol-dependent pathway. However, we did not isolate any single transpositional mutant of Tol-dependent pathway related genes that would be resistant to Kvarla. It certainly could be related to the limits of the method used, as Kvarla-resistant transposon mutants were very hard to obtain and only some false-positive clones were isolated. Mutations with high fitness penalty might not have been obtained in the conditions used for selection. Thus, we could thus far identify only one protein participating in

Kvarla reception–translocation, the outer membrane protein C (OmpC). OmpC mutants were selected by their resistance to KpneA and KaerA and it appeared that they are equally resistant to Kvarla.

The further elucidation of klebicin receptors and translocators is important also for practical use of these klebicins. Klebicins are most promising for use for fighting antibiotic-resistant strains. However, it has been shown that 97.1 % of carbapenem-resistant Klebsiella strains do not express or express less OmpC or OmpF (Ye et al., 2018). We did not test carbapenem–resistant strains in our study, but it indicates that carbapenem-resistant Klebsiella could be expected to be resistant to KpneA, KaerA and Kvarla, as all these klebicins require functional OmpC for their activity. The next step would be the attempt to change the specificity of klebicins by engineering the proteins, for example by swapping their receptor-translocation and killing domains.

Meanwhile, we can conclude that in the current state of research we have a panel of six highly efficient plant–expressed klebicins, which can together target about 91% of tested clinical strains. Even without further engineering and improvement, these proteins could be further developed for their potential use in medicine as antimicrobials against antibiotic-resistant Klebsiella.

**Example 10: MIC determination for klebicins KpneA, KaerA, KpneM, KpneM2, KvarM and Kvarla against several *Klebsiella* species**

The minimum inhibitory concentration (MIC) was calculated as the lowest concentration of bacteriocin, which prevented visible growth of corresponding bacterial strain. For determination of MICs of individual klebicins, purified lyophilized KpneA, KaerA, KpneM, KpneM2 and KvarM proteins were dissolved in sterile distilled water at the concentration of 0.5 µg/µl. For each individual bacteriocin, serial 2-fold dilutions in MHB medium (Müller-Hinton Broth; Mueller & Hinton (1941) Experimental Biology and Medicine. 48(1): 330–333) were prepared. 10 µl aliquots of each protein dilution were loaded into empty wells of sterile 96 well microplates. Two repeats of every evaluation point were made.

Overnight bacterial cultures grown in MHB medium were diluted to $OD_{595}$=0.5 in 1 ml of MHB, then diluted 1000-fold in 10 ml of the same medium. 90 µl aliquots of diluted bacterial suspensions were loaded using multichannel pipette into each well of 96-well microplate already containing protein dilutions. Additional aliquots of diluted bacterial suspension were plated on MHA medium (Müller-Hinton Agar; MHB containing 1.7% agar) for CFU enumeration in initial bacterial inoculum. For bacterial growth, microplates and agar plates were incubated at 30°C or 37°C depending on the optimal growth conditions for *Klebsiella* strains for 20h. *Klebsiella pneumoniae subsp. ozaenae* DSM 16358, *Klebsiella variicola* DSM 15968 and *Klebsiella*

*aerogenes* DSM 30053 were incubated at 30°C, *Klebsiella quasipneumoniae subsp. similipneumoniae* DSM 28212 and *Klebsiella oxytoca* DSM 5175 at 37°C.

MICs were determined by visual examination of bacterial growth in microplate wells. If the difference between two repeats was in one protein dilution, the MIC was determined as a mean of two concentrations.

Table 7 shows MIC values of 6 klebicins against 5 selected sensitive *Klebsiella* strains determined in µg protein/ml solution as well as in nM (µM).

**Table 7** MICs of klebicin bacteriocins against selected *Klebsiella* strains.

| Strain / Klebicin | *Klebsiella pneumoniae subsp. ozaenae* DSM 16358 | *Klebsiella quasipneumoniae subsp. similipneumoniae* DSM 28212 | *Klebsiella oxytoca* DSM 5175 (ATCC 13182) | *Klebsiella variicola* DSM 15968 (ATCC BAA-830) | *Klebsiella aerogenes* DSM 30053 |
|---|---|---|---|---|---|
| KpneA | >50 µg/ml<br>>1.3 µM | 0.6 µg/ml<br>15 nM | NT | 0.2 µg/ml<br>3.8 nM | NT |
| KaerA | >50 µg/ml<br>>1.3 µM | 0.2 µg/ml<br>5.0 nM | NT | 0.1 µg/ml<br>2.5 nM | NT |
| KpneM | 0.6 µg/ml<br>19.5 nM | 1.9 µg/ml<br>63 nM | NT | 0.39 µg/ml<br>13 nM | 1.2 µg/ml<br>40 nM |
| KpneM2 | 0.8 µg/ml<br>27 nM | >50 µg/ml<br>>1.7 µM | NT | NT | 0.2 µg/ml<br>6.7 nM |
| KvarM | 0.1 µg/ml<br>3.3 nM | >50 µg/ml<br>>1.7 µM | 37.5 µg/ml<br>1.3 µM | 0.2 µg/ml<br>6.6 nM | 0.6 µg/ml<br>20 nM |
| Kvarla | NT | 1.2 µg/ml<br>27.5 nM | 2.4 µg/ml<br>54 nM | 0.39 µg/mL<br>9.0 nM | 0.1 µg/mL<br>2.3 nM |

In most cases, determined klebicin MIC values were below 1-2 µg/ml. Nguen et al. (Scientific Reports (2018) 8: 241) determined MICs of 20 conventional antibiotics against 1497 strains of *Klebsiella*. Typically, MIC values determined in this study were found between 0.5 and 32 µg/ml. Thus, klebicins are comparable or superior to conventional antibiotics in terms of antibacterial activity calculated on the weight basis. Given the difference in molecular weight (most antibiotics have MW of less than 1 kDa), klebicins have significantly higher antibacterial activity calculated on the molar basis.

**Example 11: Stability of klebicins KpneA, KaerA, KpneM, KpneM2, KvarM and Kvarla upon the storage**

For stability evaluation, purified lyophilized klebicin protein samples were stored at -20°C, 5°C and room temperature (approx. 23 °C). Protein stability was assessed on the basis of antimicrobial activity at following time points: the day 0, 1 week, 2 weeks, 3 weeks, 5 weeks, 3 months, 6 months, 10 months and 12 months of storage. Protein activity against susceptible bacterium was evaluated in liquid cultures or by radial diffusion assay. Klebicins were tested with next strains: KpneM and KpneM2 with *K. pneumoniae* DSM16358, Kvarla with *K. oxytoca* DSM5175, KpneA, KaerA and KvarM with *K. quasipneumoniae* DSM28212.

Lyophilized protein samples were resuspended in distilled water (0.2-0.4 mg/ml). Soluble protein concentration was measured for each sample using Bradford assay. For the stability evaluation in liquid culture, 5 µg of bacteriocin solution were added to 1 ml of the suspension of susceptible bacterial strain of $OD_{600}$=0.3 in CAA medium ("0" time point). Bacteria mixed with bacteriocin were incubated for 4.5 h at the shaker. *K. oxytoca* DSM5175 and *K. quasipneumoniae* DSM28212 were incubated at 37°C, *K. pneumoniae* DSM16358 was incubated at 30°C. Serial dilutions were made in LB medium. Bacteria were plated on LB agar plates and incubated overnight at 30°C or 37°C, then CFU was calculated. Antimicrobial activity was evaluated as CFU/mL $\Delta\log_{10}$ in regard to untreated sample.

For radial diffusion assay, serial 1:2 dilutions of proteins solubilized in PBS buffer were made. 5 µL of protein dilutions (1-1.8 µg of protein before dilution), were spotted on soft agar plates with susceptible bacterial strain. Residual activity of klebicins was evaluated after o/n incubation of plates. Antimicrobial activity was evaluated as specific activity units (AU) - highest dilution giving a difference to non-affected bacterial growth area determined by visual inspection of plates for bacterial growth inhibition by holding the plate in front of a light source. Highest dilution with growth inhibition was recorded as an activity in AU/µg of bacteriocins. All experiments were performed in triplicate.

At -20 °C, activity of all the klebicins KpneM, KpneM2, KvarM, KpneA, KaerA and Kvarla remained stable throughout the whole year (Fig. 11A,D), despite the concentrations of soluble proteins decreased insignificantly (Fig. 11G).

Upon the storage at 5°C, five out of six klebicins remained active for one year; the activity of KpneA decreased (Fig. 11E). Concentrations of KvarM and KpneM2 in solution decreased significantly suggesting certain drop in solubility upon the storage (Fig. 11H).

In general, klebicins were less stable during storage at room temperature. Nevertheless, activities of klebicins Kvarla and KpneM remained stable for one year. Activities of KpneA, KvarM and KpneM2 decreased dramatically (Fig. 11C). This reduction in activity correlates with the concentration of protein in solution suggesting decrease in solubility upon the storage (Fig. 11I).

**Example 12: Evaluation of bacteriocin Kvarla activity against *Klebsiella quasipneumoniae* in mice gastrointestinal model**

*K. pneumoniae* is the normal component of human gut microbiota. Gastrointestinal carriage has been regarded as a major reservoir of *K. pneumoniae* infections, especially in intensive care patients (Gorrie et al., 2017). A prospective study in 1971 indicated that 18.5% patients colonized with multidrug-resistant *K. pneumoniae* after hospital admission had higher risk to develop subsequent infection caused by identical bacteria within 21 days compared to those who did not become intestinal carriers (45% vs. 11%) (Martin and Bachman, 2018).

Orally delivered klebicins could be an efficient tool for the eradication of symptomless multiresistant *K. pneumoniae* from the gut of hospitalized patients. As klebicins are quickly inactivated by gastrointestinal enzymes in case of oral admission, they should be formulated for gastric protection and release in the small and large intestine. In this example, klebicin Kvarla was formulated with Eudragit S100 for ileum and colon delivery (release of klebicin at pH above 7) and administered by oral gavage to mice with *K. quasipneumoniae* colonized gut.

Coating of KvarlA

5% Eudragit S100 solution was prepared by dissolving 0,5 g Eudragit S100 (Evonik) in 10 ml of miliQ $H_2O$ and by sonication in ultrasonic bath for 30 min at 25°C. 250 µg of Kvarla was dissolved in 200 µg of 5% Eudragit S100. Resulting solution was lyophilized at -51°C for 24 h.

Simulated gastric digestion, activity evaluation by radial diffusion assay.

To find out if Eudragit-S100-coated Kvarla is resistant to pepsin digestion, simulated gastric digestion experiment was performed. Exposures of the proteins to Simulated Gastric Fluid (SGF, commercial acidic pepsin extract) were done using low enzyme-to-substrate ratios. Methods were derived from Moreno et al. (2005), Mandalari et al. (2009) and Eiwegger et al. (2006). Briefly, plant-produced Kvarla and Eudragit-S100-coated Kvarla were mixed with SGF in recommended concentration and incubated for up to 60 min at 37°C. Every few minutes, samples of digestion mix were taken for analysis; the digestion of the protein into fragments was assessed using SDS-PAGE; in parallel, residual antimicrobial activity was evaluated using radial diffusion assay. Coomassie staining on gels was used to visualize protein decomposition and estimate the MW of peptide products, but this method was only used for uncoated Kvarla, as Eudragit S100 distorted protein migration on the SDS-PAGE gel.

Protein samples were incubated at 37°C, with rotation at 200 rpm for 10 min. Pepsin (0.15 M NaCl, 5 mg/ml) was added to give 80-113 U (pepsin:protein ratio 1:40) of pepsin per mg of protein in the final digestion mix containing 1 mg of protein and 0,025 mg of pepsin. The samples

were placed in shaker (200 rpm, 37 °C). Aliquots of reaction (50 µl) was removed at different time points (0.5, 5, 10, 20, 30 and 60 min). Digestions were stopped by raising the pH to 6.5 by addition of 0.5 M ammonium bicarbonate (10 µl of $NH_4HCO_3$) to inactivate pepsin.

Eudragit-coated Kvarla samples were adjusted to pH 8 to get Eudragit coat dissolved. Then dilutions of all samples by ratio 1:2 were made in distilled water and 5 µL aliquots of diluted samples were dropped on MHA plates with *K.quasipneumoniae* DSM28212 lawn for soft agar overlay assay.

It appears that under the conditions used (pepsin:protein ratio 1:40), protein coating with Eudragit S100 is able to provide temporal resistance to pepsin digestion. Coated Kvarla demonstrated still detectable activity in agar diffusion assay after 20 min of *in vitro* gastric digestion, while non-coated Kvarla was inactivated in simulated gastric juice very quickly, and completely lost its activity already after 0.5 min of digestion (Fig.12A). From SDS-PAGE profile of uncoated Kvarla digestion products, it is apparent that uncoated Kvarla is digested by pepsin very rapidly, and the full length protein in undetectable on the gel after 5 min of digestion (Fig. 12B).

Colonization of mice gut by *K. quasipneumoniae* DSM28212 and Kvarla treatment

Before the experiment, BALB/c mice (n=12) were acclimated in individual cages for 3 days. Three consecutive days, mice were given drinking water with ampicillin (2000 U/ml) and streptomycin (2 mg/ml) for eradication of gut microflora. Ampicillin in drinking water was continued till the end of experiment.

At $4^{th}$-to-$6^{th}$ and $11^{th}$-to-$13^{th}$ days of experiment, the mice were given using gavage $10^9$ cfu of *K. quasipneumoniae* DSM28212 once daily. Five days after last *K. quasipneumoniae* gavage ($18^{th}$ day) mice were separated into four groups (n=3): first group was given PBS gavage, second group was given 100 µg of Kvarla, third group was given 100 µg of Eudragit-S100-coated Kvarla, and fourth group was given 1000 µg of Eudragit-S100-coated Kvarla. The gavage was continued $18^{th}$-to-$21^{nd}$ day (four days), once daily. The faecal samples were collected before the inoculation of *K. quasipneumoniae*, then daily from $18^{th}$ (just before the start of the Kvarla treatment) to $22^{nd}$ day of experiment.

During therapy, the cages were replaced daily. The mice received normal diet and food and water *ad libitum* during all the experiment, but 6 hours before the gavage the animals were starved.

Faeces collected at $18^{th}$ day of experiment (just before the start of Kvarla treatment) and $22^{nd}$ day of experiment (one day after last Kvarla treatment) were used to quantify the amount of *K. quasipneumoniae* DNA by real time-PCR.

Real time PCR

DNA was extracted from 50 mg of faeces by use of QIAamp Fast DNA Stool Mini Kit (Qiagen). *Klebsiella* hemolysin gene (*khe*) marker was used for amplification. *khe* gene amplification primers used: Forward: 5 -GATGAAACGACCTGATTGCATTC-3 (SEQ ID NO: 56), Reverse: 5 -CCGGGCTGTCGGGATAAG-3 (SEQ ID NO: 57), Probe: 5 -6FAM-CGCGAACTGGAAGGGCCCG-TAMRA-3 (SEQ ID NO: 58). „TaqMan Universal Master Mix II with UNG" and „TaqMan probe" (Applied Biosystems, JAV) were used. 14 ng of DNA was used for each PCR reaction. Following controls were used for real time PCR: *K. quasipneumoniae* DSM28212 DNA (*Klebsiella* hemolysin gene *khe* amplicifation in cycle 13), *E. coli* DNA – no *khe* amplification and blanc – no *khe* amplification.

**Table 8** Real time-PCR conditions.

| | UNG incubation | Activation of polymerase | PCR (40 cycles) | |
|---|---|---|---|---|
| | | | Denaturation | Elongation |
| Temperature (C °) | 50 | 95 | 95 | 60 |
| Time | 0:20 | 10:00 | 0:15 | 1:00 |

Real time PCR was validated by setting up a standard curve for detection of *K. quasipneumoniae*. According to this curve, CT 38 correspond to $10^3$ CFU, CT 24 corresponds to $10^6$ CFU and CT 17- to $10^8$ CFU of *K. quasipneumoniae* (Fig. 13).

Colonization of mice gut by *K. quasipneumoniae* DSM28212 before and after klebicin gavage

Real time PCR results confirmed that at 18$^{th}$ day of experiment (before the start of klebicin treatment) all mice demonstrated the presence of *K. quasipneumoniae* DNA in faeces. At 18$^{th}$ day of experiment, median CT values of each group of mice were in range of 18.3-20.5 (Fig. 14 and Table 9).

According to real time PCR results, the numbers of *K. quasipneumoniae* kept rising in the faeces of groups of mice treated by PBS and by uncoated Kvarla. At the day 22, the day after last klebicin treatment, median CT value in PBS group decreased from 19.97 to 17.23, and in Kvarla treated group from 20.05 to 18.36.

In contrast to the PBS-treated and uncoated Kvarla-treated mice, the amount of *K. quasipneumoniae* DNA sharply decreased in faeces of mice treated by Eudragit-S100-coated Kvarla (both concentrations). Median CT values increased by 8 – 8.5 cycles (from 18.5 to 26.3 for Eudragit_S100-Kvarla 100 µg group and from 18.86 to 27.43 for Eudragit_S100-Kvarla 1 mg group) (Fig. 14 and Table 9).

**Table 9** CT values of each individual mouse after *Klebsiella* colonization and after therapy.

| **Day 18 (before treatment)** | | | **Day 22 (after treatment)** | |
|---|---|---|---|---|
| **GROUP 1, PBS** | **CT value** | | **GROUP 1, PBS** | **CT value** |
| Mouse 1 | 17.79 | | Mouse 1 | 15.98 |
| Mouse 2 | 20.59 | | Mouse 2 | 15.00 |
| Mouse 3 | 21.53 | | Mouse 3 | 20.72 |
| **Mean** | **19.97** | | **Mean** | **17.23** |
| **GROUP 2, day 18 Uncoated Kvarla 100 µg** | **CT value** | | **GROUP 2, day 22 Uncoated Kvarla 100 µg** | **CT value** |
| Mouse 1 | 17.31 | | Mouse 1 | 17.34 |
| Mouse 2 | 20.13 | | Mouse 2 | 16.38 |
| Mouse 3 | 22.70 | | Mouse 3 | 21.38 |
| **Mean** | **20.05** | | **Mean** | **18.36** |
| **GROUP 3, day 18 Eudragit_S100-Kvarla 100 µg** | **CT value** | | **GROUP 3, day 22 Eudragit_S100-Kvarla 100 µg** | **CT value** |
| Mouse 1 | 20.68 | | Mouse 1 | 26.53 |
| Mouse 2 | 16.82 | | Mouse 2 | 25.08 |
| Mouse 3 | 17.41 | | Mouse 3 | 27.29 |
| **Mean** | **18.30** | | **Mean** | **26.30** |
| **GROUP 4, day 18 Eudragit_S100-Kvarla 1 mg** | **CT value** | | **GROUP 4, day 22 Eudragit_S100-Kvarla 1 mg** | **CT value** |
| Mouse 1 | 17.28 | | Mouse 1 | 27.50 |
| Mouse 2 | 18.57 | | Mouse 2 | 29.08 |
| Mouse 3 | 20.72 | | Mouse 3 | 25.71 |
| **Mean** | **18.86** | | **Mean** | **27.43** |

Thus, according to real time PCR results, Eudragit-S100-coated Kvarla drastically decreased the *K. quasipneumoniae* DNA amount in the faeces of mice. The obtained decrease was similar for both dosages used: 100 µg and 1 mg. By contrast, the *K. quasipneumoniae* DNA amount slightly increased in uncoated Kvarla-treated and PBS-treated mice.

In conclusion, Eudragit-S100-coated Kvarla demonstrated high activity in reducing *K. quasipneumoniae* DNA amount in the gut of mice, which is indicative of the decrease in the population of this bacterium.

NUCLEIC ACID AND AMINO ACID SEQUENCES

SEQ ID NO:1 referred to as KpneM (*K. pneumoniae* EWD35590.1)
MSETMVVVATPTGFEPAGYGGGLFSPSTPNHSPSQGQIFLQVTLPYYQSTKFCQDSMAWLAQYVKTHGA
QDPLTIQVVANNIRYFLNADTNLCHNPKQNVWEAFHSEMTHSGPPPAKYDYHSMSLKQMSGN
VVTPAAA
FGHYLWGNGEARYVNLPDVGLKITPQMIPELMNIVNSGVTGHIPVDIKFVHDTSVSGGIVPAAYLGHITLRT
EGTLDIQSGGAWTYNGVARAF NDTYDFNLGDFRGPIAESMTFLGSQFTGKQYEISMPGQINISGSGRR

SEQ ID NO:2 referred to as KvarM (*K. variicola* CTQ17225.1)
MSDTMIVVATPTPGFSYASGLTYGGGAFAGAPANGPSEGQIFFQTVLPAYQSPNLCIGQLAWMTDYINKN
GVGNPKTWEVISQNVLIFCSADTALVLNPRIAVYDGFHKTKWAPAKFNFKTQSQEKFSGNVTTPIAAFG
HYLWGEGKPRTVDLSSVGLKIQANQIDPVMIAVKNNAAGTYQISGNFNRNTFIDGDIPGLYLGNITMKTEGT
LKIDAKGNWNYNGVVRAFNDTYDANPSTHRSKSAEDLTTLLRLTQGTPYEIRIPGELKVSGSGKK

SEQ ID NO:3 referred to as KpneM2 (*Klebsiella sp.* WP_047066220)
MSETLVVVAPAPSAPSMTYGGGLIYSSIPSGPNEGQIFFQTVLPAYSSPNFCTDRLRWMVKFINENGVGNP
DTWKTLADVIRYYASADTAISKNPKTNPYDAWHKCPWPPASFDVKTMSVEKFSGSVNTPIVAFGHYL
WGEGKPRSVDLSTVGLKVQANQIDPVMIAVKSYGAGTYQINGNFNRNTFDDGVIPGLYLGNITLKTEGTLKI
EKNGSWNYNGVIRAFNDTYDANPSN HRSQAAEDLTTLLRITQGTPYEIRIPGEIKVSGSGKK

SEQ ID NO:4 referred to as KaerM (*K. aerogenes* WP_015367360.1)
MTDTLTVTATIPNGSSFNFQFEGMGNYYAAGSSTWDDPAMADAAHLYNAIQSMEDGSFTKALFADWLQF
NAKGRENIPMINARFATMETMRFNDPGKAYFQFAQYNEYEGHTPGNNFTSGAFAPFLGLWHYISGNGVE
TSLDITTIGLTFNQSNLTPVNDALKSQPPGNYPISSNFGKSVAEDNLYVAALLGRISMKTEGTLSIGESGEWSY
NGVVRAYNDTYDANFDPSRGVIAQASTTVLSWFNGKPYPIALPGEIPVQLSGHR

SEQ ID NO:5 referred to as KpneA (*K. pneumoniae* SAV78255.1)
MPEETLTVVGGGNNSCNVSWGGGNGNNGGAGYSGKYGGTSYEGATSMLKLNDRVLIQLYLCNPLNPDYI
GAPWGSDKDAESIIRANRDKPGKFKANIQNWKTSGTGSLGSPVVGKSYSSGDVDTYSVSFGKEKYNVLYNR
KKDSFTTAYVDGGANKPEHSMKDQAIAVVKLYLLNESQASVIDTTSGIITDSGKTLSGKLGDKYNTLAREAAD
NIKNFQGKKLRSFNDAMASINELANNPKMKLSQADKTVVSNALKQMDLSALADRFKGLEKAFTWGDRLLK
AEKIRDGVVTGVTTGDWQKLAFEVEAMYLSGVAGAVALGITTAMISTVAVALSLPSVAVSALTVVAVIGI
SILTSYIDADKAKALNNAVLGLFK

SEQ ID NO:6 referred to as KaerA (*K. aerogenes* WP_063414841.1)
MANEDSMTVNGNAGSGVHWGGGSGNGNNGGAGSNGGANVALGGTMEVELGNGFTMIVDGTHPINP
GIGGAPWSDDKSNKSAVDALNANKSKPAKFKANIQNYKSGTQGSLNSPAVNKSSSSGDVDTYAVSFGKEKY

NVMYNRKKDSFTSGYVDGGATKPEHSMKDQAIAVVQLYLLNEKEKDVITTAAEIISSSGETISGKLGEKYKGL
AQGVANDIRNFQGKKIRSFKDAMSSLEQFTKNPNMKLNQADKAALVNALNQVNLSTLADRFKGLERA
FTWADRLLKAQKIKDGVVTGVTTGNWQPLALEVEAMYLSGVAGSVALGIVTGMISGLAALISIPALAVTAL
TVTAVIGIAIATSYINADTAKALNNAVADLFK

SEQ ID NO:7 referred to as Koxy (*K. oxytoca* WP_024273778)
MAGFSYGGFGDGTTWSKERGTGPLPGGGSSGNSGNHSNTTPAEQKQINAIRADKNVRARLSNLIKAARKL
NPSVKITVHAISPEGTMAISMEGLTATQARQAGLTGLVMGITVPGYIGSVGDFETGHKYNLKNPEKLNS
IGVGTPLDGFNGGENIDTTPKKYRNWRATDEKSFYYVGTTVPMRLLHHLTVSRNKETDTYTMYFKAKDIKAL
YKIEVKNGDLDNMKLTTLAQGHPLFTAEFAKDIVRNFASVKNESDKEVLDKTSGVIISVGDKAGALLGEKY
KALSREVASNIQNFQGKQIRTYDQAMASMNKLMTNPNMKIKAADKTAVINAWKAFNVEDMGNKFTALG
RAFKVADYVTKGNNVREKSITGYETGNWGPLMREVESWTVSGLTSSVALAVFSATLGAMLVAAGVST
AVVGIIGIIIAGLIGALIDDKFIDKLNNEIIRPAY

SEQ ID NO:8 referred to as Kpnela (*K. pneumoniae* BAS34675)
MPGFNYGGKGDGTNWSSERGTGPEPGGGSRGNGGDRDNSRGGAGNRGNWAGSGPLSAALINDSIAEAL
EKQLPRNTVEATSTPAYKKMRAAFDALPLDKQPEARAQITKAWQSAHDAMPDKTTTTENVGGGKN
GHNVTRSTPNWLKEKMKGLNQQVNNDLSGALAQHQKAEADARAKAEAAAKAKAEAEAKAKAEAEAKAK
AKAEAAAKAKAEAEAKAKAEAEAKAKAEAEAAAKAKAEAEAKAKAEAEAKAKAEAEAAAKAKAEAEAKAKAE
AEAKAKAEAEAKAKAEADAVKDAVKFTADFYKEVFSVYGEKAEQLANLLATQAKGKNIRNIDDALKAYEKHK
TNINKKINAQDRAAIAKALESVDVKEAAKNFAKFSKGLGYVGPTMDVVDLVLELRKAIKEDNWRSFFV
KIEAIAISFGATQLAALAFASLLGAPVGLLGYALIMAGIGALVSDDVVDAANKIIGI

SEQ ID NO:9 referred to as Kvarla (*K. variicola* KDL88409)
MPGFNYGGKGDGTNWSSERGTGPEPGGGSRGNGGDRDNSRGGAGNRGNWAGSGPLSAALINDSIAEAL
EKQLPRNTVEATSTPAYKKMRAAFDALPLDKQPEARAQITKAWQSAHDAMPDRTTTTENVGGGKNGH
NVTRSTPNWLKEKMKGLNQQVNNDLSGALAQHQKAEADARAKAEAAAKAKAAAKAKAEAEAKAKAEAE
AKAKAEAAAKAKAEAEAKAKAEAEAKAKAEADAVKDAVKFTADFYKEVFSVYGEKAEQLANLLAT
QAKGKNIR NIDDALKAYEKHKTNINKKINAQDRAAIAKALESVDVKEAAKNFAKFSKGL
GYVGPTMDVVDLVLELRKAIKEDNWRTFFVKIEAIAISFGATQLAALAFASLLGAPVGLLG
YALIMAGIGALVSDDVVDAANKIIGI

SEQ ID NO: 10: ExbB Eco88I fwd
AAACTCGGGTTGATGAACCTGTTTTTATACGTCT

SEQ ID NO: 11 ExbB Eco81I rev
AAACCTGAGGTCAACCTACCCGTAATTTCTGCG

SEQ ID NO: 12 ExbB Eco88I fwd
AAACTCGGGTTGATGAACCTGTTTTTATACGTCT

SEQ ID NO: 13 ExbD Eco81I rev
AAACCTGAGGTTATTTGGCTTTGACGGTCTC

SEQ ID NO: 14 FhuA Eco81I fwd
AAACCTCAGGTTTAAGCCCTAAGACCAGACCC

SEQ ID NO: 15 FhuA Eco 81I rev
AAACCTGAGGTTAGAAACGGAAGGTGGCGGTG

SEQ ID NO: 16 FimB Eco88I fwd
AAACTCGGGGCTCCCGTAGCAAATAAAAACG

SEQ ID NO: 17 FimB Eco81I rev
AAACCTGAGGTTACTGAAGCAGCGACAGGCG

SEQ ID NO: 18 OmpC Eco88I fwd
AAACTCGGGCTTGTGGCTGAACGACTCATCA

SEQ ID NO: 19 OmpC Eco81I rev
AAACCTGAGGTTAGAACTGGTAAACCAGGCCC

SEQ ID NO: 20 TonB PsyI fwd
AAAGACCGGGTCGGCAAAGCTCCTTATCAATAAACA

SEQ ID NO: 21 TonB BseSI rev
AAAGTGCCCTCAGTTAATCTCGACGCCGTTG

SEQ ID NO: 22 Consensus sequence M (KpneM2, KvarM, KpneM, KaerM)
MSXTXVVVATPXXXXXXXXTYGGGLFYXXXPXGPSEGQIFFQTVLPAYQSPNFCXDXLAWMA
DYINXNGVGNPXTWEVIAXNIRYFASADTALXXNPKXXVYDAFHKXXWPPAKXDXXTMSXEKF
SGNVXTPIAAFGHYLWGXGKPRSVDLSTVGLKIQANQIDPVMIAVKSXXAGTYXISGNFNRNTF
XDGXIPXXYLGNITXKTEGTLKIXXXGXWNYNGVVRAFNDTYDANPSXHRXXIAEDLTTLLXXX
QGXPYEIRIPGEIKVSGSGKX

SEQ ID NO: 23 Consensus sequence A (KaerA, KpneA)

MXXEXXXXVXGXNXXXXVXWGGXXGNGNNGGAGXXGXXGXXXXXGXTXXXXLXBXXXXXX
XXXXPJNPXXXGAPWXXXXSBKXAXXXJXANXXKPXKFKANIQNXKXXXXGSLXSPXVXKSXS
SGDVDTYXVSFGKEKYNVXYNRKKDSFTXXYVDGGAXKPEHSMKDQAIAVVXLYLLNEXZXX
VIXTXXXIIXXSGXTJSGKLGXKYXXLAXXXABBIXNFQGKKJRSFXDAMXSJXZXXXNPXMKLX
QADKXXXXNALXQXBLSXLADRFKGLEXAFTWXDRLLKAZKIXDGVVTGVTTGBWQXLAXEVE
AMYLSGVAGXVALGIXTXMISXXAXXJSJPXXAVXALTVXAVIGIXIXTSYIBADXAKALNNAVXXL
FK

SEQ ID NO: 24 Consensus sequence Ia (Kpnela, Kvarla)
MPGFNYGGKGDGTNWSSERGTGPEPGGGSRGNGGDRDNSRGGAGNRGNWAGSGPLSAA
LINDSIAEALEKQLPRNTVEATSTPAYKKMRAAFDALPLDKQPEARAQITKAWQSAHDAMPDX
TTTTENVGGGKNGHNVTRSTPNWLKEKMKGLNQQVNNDLSGALAQHQKAEADARAKAEAAA
KAKXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXAXAKAKAE
AEAKAKAEAXAKAKAEAXAKAKAEAEAKAKAEAEAKAKAEADAVKDAVKFTADFYKEVFSVYG
EKAEQLANLLATQAKGKNIRNIDDALKAYEKHKTNINKKINAQDRAAIAKALESVDVKEAAKNFA
KFSKGLGYVGPTMDVVDLVLELRKAIKEDNWRXFFVKIEAIAISFGATQLAALAFASLLGAPVGL
LGYALIMAGIGALVSDDVVDAANKIIGI

SEQ ID NO: 25 conserved sequence stretch in SEQ ID NO: 22
TEGTL

SEQ ID NO: 26 conserved sequence stretch in SEQ ID NO: 22
YNGV

SEQ ID NO: 27 conserved sequence stretch in SEQ ID NO: 22
RAFNDTYD

SEQ ID NO: 28 conserved sequence stretch in SEQ ID NO: 23
GNGNNGGAG

SEQ ID NO: 29 conserved sequence stretch in SEQ ID NO: 23
GAPW

SEQ ID NO: 30 conserved sequence stretch in SEQ ID NO: 23
KFKANIQN

SEQ ID NO: 31 conserved sequence stretch in SEQ ID NO: 23
SSGDVDTY

SEQ ID NO: 32 conserved sequence stretch in SEQ ID NO: 23
VSFGKEKYNV

SEQ ID NO: 33 conserved sequence stretch in SEQ ID NO: 23
YNRKKDSFT

SEQ ID NO: 34 conserved sequence stretch in SEQ ID NO: 23
YVDGGA

SEQ ID NO: 35 conserved sequence stretch in SEQ ID NO: 23
KPEHSMKDQAIAVV

SEQ ID NO: 36 conserved sequence stretch in SEQ ID NO: 23
LYLLNE

SEQ ID NO: 37 conserved sequence stretch in SEQ ID NO: 23
SGKLG

SEQ ID NO: 38 conserved sequence stretch in SEQ ID NO: 23
NFQGKK

SEQ ID NO: 39 conserved sequence stretch in SEQ ID NO: 23
QADK

SEQ ID NO: 40 conserved sequence stretch in SEQ ID NO: 23
LADRFKGL

SEQ ID NO: 41 conserved sequence stretch in SEQ ID NO: 23
AFTW

SEQ ID NO: 42 conserved sequence stretch in SEQ ID NO: 23
DRLLKA

SEQ ID NO: 43 conserved sequence stretch in SEQ ID NO: 23
DGVVTGVTTG

SEQ ID NO: 44 conserved sequence stretch in SEQ ID NO: 23

EVEAMYLSGVAG

SEQ ID NO: 45 conserved sequence stretch in SEQ ID NO: 23
VALGI

SEQ ID NO: 46 conserved sequence stretch in SEQ ID NO: 23
ALTV

SEQ ID NO: 47 conserved sequence stretch in SEQ ID NO: 23
AVIGI

SEQ ID NO: 48 conserved sequence stretch in SEQ ID NO: 23
TSYI

SEQ ID NO: 49 conserved sequence stretch in SEQ ID NO: 23
AKALNNAV

SEQ ID NO: 50 conserved sequence stretch in SEQ ID NO: 24
MPGFNYGGKGDGTNWSSERGTGPEPGGGSRGNGGDRDNSRGGAGNRGNWAGSGPLSAA
LINDSIAEALEKQLPRNTVEATSTPAYKKMRAAFDALPLDKQPEARAQITKAWQSAHDAMPD

SEQ ID NO: 51 conserved sequence stretch in SEQ ID NO: 24
TTTTENVGGGKNGHNVTRSTPNWLKEKMKGLNQQVNNDLSGALAQHQKAEADARAKAEAAA
KAK

SEQ ID NO: 52 conserved sequence stretch in SEQ ID NO: 24
AKAKAEAEAKAKAEA

SEQ ID NO: 53 conserved sequence stretch in SEQ ID NO: 24
AKAKAEA

SEQ ID NO: 54 conserved sequence stretch in SEQ ID NO: 24
AKAKAEAEAKAKAEAEAKAKAEADAVKDAVKFTADFYKEVFSVYGEKAEQLANLLATQAKGKN
IRNIDDALKAYEKHKTNINKKINAQDRAAIAKALESVDVKEAAKNFAKFSKGLGYVGPTMDVVDL
VLELRKAIKEDNWR

SEQ ID NO: 55 conserved sequence stretch in SEQ ID NO: 24
FFVKIEAIAISFGATQLAALAFASLLGAPVGLLGYALIMAGIGALVSDDVVDAANKIIGI

SEQ ID NO: 56 *khe* gene amplification forward primer
GATGAAACGACCTGATTGCATTC

SEQ ID NO: 57 *khe* gene amplification reverse primer
CCGGGCTGTCGGGATAAG

SEQ ID NO: 58 *khe* gene amplification probe sequence, labelled with 6FAM at the 5'-end and with TAMRA at the 3'-end
CGCGAACTGGAAGGGCCCG

REFERENCES

Bodier-Montagutelli E, Mayora A, Vecellioa L, Respauda R, Heuzé-Vourc'h N. Designing inhaled protein therapeutics for topical lung delivery: what are the next steps? Expert Opin Drug Deliv. 2018: 15 (8): 729–736

Bulet P, Menin SRL. Anti-microbial peptides: from invertebrates to vertebrates. Immunol Rev 2004; 198:169–84

Chan YR, Liu JS, Pociask DA, Zheng M, Mietzner TA, Berger T, et al. 2009. Lipocalin 2 is required for pulmonary host defense against *Klebsiella* infection. J Immunol. Apr 15 182(8):4947-56.

Chavan M1, Rafi H, Wertz J, Goldstone C, Riley MA. Phage associated bacteriocins reveal a novel mechanism for bacteriocin diversification in *Klebsiella*. J Mol Evol. 2005 Apr;60(4):546-56)

Depreter F, Pilcer G, Amighi K. Inhaled proteins: challenges and perspectives. Int J Pharm. 2013: 447(1-2):251-80.

Eiwegger T, Rigbyw N, Mondouletz L, Bernardz H, Krauth M.T, Boehm A, Dehlink E, Valent P, Walz J.M, Millsw E.N.C, Szépfalusi Z. Gastro-duodenal digestion products of the major peanut allergen Ara h 1 retain an allergenic potential. 2006 Clinical and Experimental Allergy, 36, 1281–1288.

Ghequire MG, De Mot R. Ribosomally encoded antibacterial proteins and peptides from *Pseudomonas*. FEMS Microbiol Rev. 2014 Jul;38(4):523-68

Grinter R, Milner J, Walker D. Beware of proteins bearing gifts: protein antibiotics that use iron as a Trojan horse. FEMS Microbiol Lett. 2013 Jan;338(1):1-9

Gorrie, C. L., Mirceta, M., Wick, R. R., Edwards, D. J., Thomson, N. R., Strugnell, R. A., et al. (2017). Gastrointestinal carriage is a major reservoir of Klebsiella pneumoniae infection in intensive care patients. Clin. Infect. Dis. 65, 208–215. cix270

Gu D, Dong N, Zheng Z, Lin D, Huang M, Wang L, Chan EW, Shu L, Yu J, Zhang R, Chen S. A fatal outbreak of ST11 carbapenem-resistant hypervirulent *Klebsiella pneumoniae* in a Chinese hospital: a molecular epidemiological study. Lancet Infect Dis. 2017 Aug 29. pii: S1473-3099(17)30489-9.

Hirche TO, Gaut JP, Heinecke JW.2005. Myeloperoxidase plays critical roles in killing *Klebsiella pneumoniae* and inactivating neutrophil elastase: effects on host defense. J Immunol. 174(3):1557-65.

Harding CR, Schroeder GN, Collins JW, Frankel G (2013) Use of *Galleria mellonella* as a Model Organism to Study *Legionella pneumophila* Infection. J Vis Experiments 81: e50964.

Kleanthous C. (2010) Swimming against the tide: progress and challenges in our understanding of colicin translocation. Nat Rev Microbiol 8 (12): 843-848.

James R1, Schneider J, Cooper PC. Characterization of three group A klebicin plasmids: localization of their E colicin immunity genes. J Gen Microbiol. 1987 Aug;133(8):2253-62.

Mandalari G, Adel-Patient K, Barkholt V, Baro C, Bennett L, Bublin M, Gaier S, Graser G, Ladics G.S, Mierzejewska D, Vassilopoulou E, Vissers Y.M, Zuidmeer L, Rigby N.M, Salt L.J, Defernez M, Mulholland F, Mackie A.R, Wickham M.S.J, Mills E.N.C. *In vitro* digestibility of b-casein and b-lactoglobulin under simulated human gastric and duodenal conditions: A multi-laboratory evaluation. Regulatory Toxicology and Pharmacology 55 (2009) 372–381.

Marillonnet S, Thoeringer C, Kandzia R, Klimyuk V, Gleba Y. (2005) Systemic *Agrobacterium tumefaciens* mediated transfection of viral replicons for efficient transient expression in plants. Nat Biotechnol 23: 718-723.

Martin, R. M., and Bachman, M. A. (2018). Colonization, infection, and the accessory genome of *Klebsiella pneumoniae.* Front. Cell. Infect. Microbiol. 8:4.

Martínez-García E, Calles B, Arévalo-Rodríguez M, de Lorenzo V. (2011) pBAM1: an all-synthetic genetic tool for analysis and construction of complex bacterial phenotypes. BMC Microbiol 11: 38.

Moreno F.J, Mellon F.A, Wickham M.S.J, Bottrill A.R, Mills E.N.C. Stability of the major allergen Brazil nut 2S albumin (Ber e 1) to physiologically relevant in vitro gastrointestinal digestion. FEBS Journal 272 (2005) 341–352.

Moskowitz SM, Foster JM, Emerson J, Burns JL. (2004) Clinically Feasible Biofilm Susceptibility Assay for Isolates of *Pseudomonas aeruginosa* from Patients with Cystic Fibrosis Clinically Feasible Biofilm Susceptibility Assay for Isolates of *Pseudomonas aeruginosa* from Patients with Cystic Fibrosis. J Clin Microbiol 42: 1915–1922.

Paškevičius Š, Starkevič U, Misiūnas A, Vitkauskienė A, Gleba Y, Ražanskienė A. (2017) Plant-expressed pyocins for control of *Pseudomonas aeruginosa*. PLoS One 12(10): e0185782.

Witt, D. M. and Anderson, L. (1996), Dornase Alfa: A New Option in the Management of Cystic Fibrosis. Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy, 16: 40-48.

Ye Y, Xu L, Han Y, Chen Z, Liu C, Ming L. (2018) Mechanism for carbapenem resistance of clinical *Enterobacteriaceae* isolates. Exp Ther Med 15(1): 1143–1149.

The content of European patent application No. 19 178 676.3, filed on June 6, 2019 is incorporated herein by reference including description, claims, figures, and sequence listing.

## Claims

1. A protein having cytotoxic activity against *Klebsiella*, or a composition comprising said protein, for use in therapy.

2. The protein or composition for the use according to claim 1, said protein having a lipid II-cleaving activity or a pore-forming capability in a cell membrane of *Klebsiella* cells.

3. The protein or composition for the use according to claim 1 or 2, for use in a method of treating infection of a subject with *Klebsiella*, such as *Klebsiella pneumoniae*, *Klebsiella oxytoca*, *Klebsiella quasipneumoniae*, *Klebsiella aerogenes*, and/or *Klebsiella variicola*, wherein said *Klebsiella* may be antibiotic-resistant such as carbapenem-resistant.

4. A protein, or composition comprising said protein, for use in a method of treating infection of a subject with *Klebsiella*, such as *Klebsiella pneumoniae*, *Klebsiella oxytoca*, *Klebsiella quasipneumoniae*, *Klebsiella aerogenes*, and/or *Klebsiella variicola*, wherein said *Klebsiella* may be antibiotic-resistant such as carbapenem-resistant.

5. A method of treating infection with *Klebsiella* of a subject in need thereof, comprising administering to said subject a protein having cytotoxic activity against *Klebsiella* or a composition comprising said protein.

6. A method of preventing or reducing infection or contamination of an object with one or more *Klebsiella species*, comprising contacting said object with a protein having cytotoxic activity against *Klebsiella*, or with a composition comprising said protein.

7. The protein, composition or method according to any one of claims 1 to 6, said protein comprising or consisting of a first amino acid sequence segment and a second amino acid sequence segment, wherein the first segment is preferably the N-terminal segment of said protein and the second segment is the C-terminal segment of said protein.

8. The protein, composition or method according to claim 7,

(A) wherein the first segment comprises or consists of the amino acid sequence of

- (A-ii) from amino acid residue 1 to 127 of SEQ ID NO: 2 (KvarM),
- (A-i) from amino acid residue 1 to 128 of SEQ ID NO: 1 (KpneM),
- (A-iii) from amino acid residue 1 to 123 of SEQ ID NO: 3 (KpneM2),
- (A-iv) from amino acid residue 1 to 118 of SEQ ID NO: 4 (KaerM),
- (A-v) from amino acid residue 1 to 170 of SEQ ID NO: 5 (KpneA),

(A-vi) from amino acid residue 1 to 172 of SEQ ID NO: 6 (KaerA),

(A-vii) from amino acid residue 1 to 255 of SEQ ID NO: 7 (Koxy),

(A-viii) from amino acid residue 1 to 288 of SEQ ID NO: 8 (Kpnela), or

(A-ix) from amino acid residue 1 to 236 of SEQ ID NO: 9 (Kvarla);

or

(B) wherein the first segment comprises an amino acid sequence

(B-ii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 127 of SEQ ID NO: 2,

(B-i) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 128 of SEQ ID NO: 1,

(B-iii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 123 of SEQ ID NO: 3,

(B-iv) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 118 of SEQ ID NO: 4,

(B-v) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 170 of SEQ ID NO: 5,

(B-vi) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 172 of SEQ ID NO: 6,

(B-vii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 255 of SEQ ID NO: 7,

(B-viii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 288 of SEQ ID NO: 8, or

(B-ix) having at least 70% sequence identity to the amino acid sequence from amino acid residue 1 to 236 of SEQ ID NO: 9;

or

(C) wherein the first segment comprises an amino acid sequence

(C-ii) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 127 SEQ ID NO: 2,

(C-i) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 128 SEQ ID NO: 1,

(C-iii) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 123 of SEQ ID NO: 3,

(C-iv) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 118 of SEQ ID NO: 4,

(C-v) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 170 SEQ ID NO: 5,

(C-vi) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 172 of SEQ ID NO: 6,

(C-vii) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 255 SEQ ID NO: 7,

(C-viii) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 288 SEQ ID NO: 8, or

(C-ix) having from 1 to 40 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 1 to 236 SEQ ID NO: 9.

9. The protein, composition or method according to claim 8, wherein in item (B) any one of the sequence identities is at least 75%, preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, even more preferably at least 95%, and most preferably at least 97%; and/or

wherein in item (C) the number of said amino acid substitutions, additions, insertions and/or deletions is from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10, and most preferably at least 1 to 5 compared to any one of said amino acid sequences.

10. The protein, composition or method according to any one of claims 7, 8 and 9,

(D) wherein the second segment comprises or consists of the amino acid sequence of

(D-ii) from amino acid residue 128 to 276 of SEQ ID NO: 2 (KvarM),

(D-i) from amino acid residue 129 to 278 of SEQ ID NO: 1 (KpneM),

(D-iii) from amino acid residue 124 to 272 of SEQ ID NO: 3 (KpneM2),

(D-iv) from amino acid residue 119 to 266 of SEQ ID NO: 4 (KaerM),

(D-v) from amino acid residue 171 to 377 of SEQ ID NO: 5 (KpneA),

(D-vi) from amino acid residue 173 to 379 of SEQ ID NO: 6 (KaerA),

(D-vii) from amino acid residue 256 to 452 of SEQ ID NO: 7 (Koxy),

(D-viii) from amino acid residue 289 to 466 of SEQ ID NO: 8 (Kpnela), or

(D-ix) from amino acid residue 237 to 414 of SEQ ID NO: 9 (Kvarla);

or

(E) wherein the second segment comprises an amino acid sequence

(E-ii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 128 to 276 of SEQ ID NO: 2,

(E-i) having at least 70% sequence identity to the amino acid sequence from amino acid residue 129 to 278 of SEQ ID NO: 1,

(E-iii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 124 to 272 of SEQ ID NO: 3,

(E-iv) having at least 70% sequence identity to the amino acid sequence from amino acid residue 119 to 266 of SEQ ID NO: 4,

(E-v) having at least 70% sequence identity to the amino acid sequence from amino acid residue 171 to 377 of SEQ ID NO: 5,

(E-vi) having at least 70% sequence identity to the amino acid sequence from amino acid residue 173 to 379 of SEQ ID NO: 6,

(E-vii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 256 to 452 of SEQ ID NO: 7,

(E-viii) having at least 70% sequence identity to the amino acid sequence from amino acid residue 289 to 466 of SEQ ID NO: 8, or

(E-ix) having at least 70% sequence identity to the amino acid sequence from amino acid residue 237 to 414 of SEQ ID NO: 9;

or

(F) wherein the second segment comprises an amino acid sequence

(F-ii) having from 1 to 30 amino acid substitutions, additions, insertions or deletions compared to the amino acid sequence of from amino acid residue 128 to 276 of SEQ ID NO: 2,

(F-i) having from 1 to 30 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 129 to 278 of SEQ ID NO: 1,

(F-iii) having from 1 to 30 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 124 to 272 of SEQ ID NO: 3,

(F-iv) having from 1 to 30 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 119 to 266 of SEQ ID NO: 4,

(F-v) having from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 171 to 377 of SEQ ID NO: 5,

(F-vi) having from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 173 to 379 of SEQ ID NO: 6,

(F-vii) having from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 256 to 452 of SEQ ID NO: 7,

(F-viii) having from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 289 to 466 of SEQ ID NO: 8,

(F-ix) having from 1 to 35 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of from amino acid residue 237 to 414 of SEQ ID NO: 9.

11. The protein, composition or method according to claim 9 or 10, wherein said first segment is any one of items A-i to A-iv, B-i to B-iv, or C-i to C-iv and said second segment is any one of items D-i to D-iv, E-i to E-iv, or F-i to F-iv.

12. The protein, composition or method according to claim 11, wherein said first segment is any one or more of items A-i to A-iv and said second segment is any one of items D-i to D-iv, respectively; or said first segment is any one of items B-i to B-iv and said second segment is any one of items E-i to E-iv, respectively; or said first segment is any one of C-i to C-iv and said second segment is any one of items F-i to F-iv, respectively.

13. The protein, composition or method according to any one of claims 1 to 6, said protein comprising or consisting of an amino acid sequence comprising or consisting of

(a) the amino acid sequence of

(a-ii) SEQ ID NO: 2 (KvarM),

(a-i) SEQ ID NO: 1 (KpneM),

(a-iii) SEQ ID NO: 3 (KpneM2),

(a-iv) SEQ ID NO: 4 (KaerM),

(a-v) SEQ ID NO: 5 (KpneA),

(a-vi) SEQ ID NO: 6 (KaerA),

(a-vii) SEQ ID NO: 7 (Koxy),

(a-viii) SEQ ID NO: 8 (KpneIa), or

(a-ix) SEQ ID NO: 9 (KvarIa);

or

(b) an amino acid sequence

(b-ii) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 2,

(b-i) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 1,

(b-iii) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 3,

(b-iv) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 4,

(b-v) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 5,

(b-vi) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 6,

(b-vii) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 7,

(b-viii) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 8, or

(b-ix) having at least 70% sequence identity to the amino acid sequence of SEQ ID NO: 9;

or

(c) an amino acid sequence

(c-ii) having from 1 to 80 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 2,

(c-i) having from 1 to 80 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 1,

(c-iii) having from 1 to 80 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 3,

(c-iv) having from 1 to 80 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 4,

(c-v) having from 1 to 110 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 5,

(c-vi) having from 1 to 110 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 6,

(c-vii) having from 1 to 130 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 7,

(c-viii) having from 1 to 130 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 8, or

(c-ix) having from 1 to 120 amino acid substitutions, additions, insertions and/or deletions compared to the amino acid sequence of SEQ ID NO: 9.

14. The protein, composition or method according to claim 13, wherein in item (b) any one of the sequence identities is at least 75%, preferably at least 80%, more preferably at least

85%, even more preferably at least 90%, even more preferably at least 95%, and most preferably at least 97%; and/or
wherein in item (c) the number of said amino acid substitutions, additions, insertions and/or deletions is from 1 to 30, preferably from 1 to 20, more preferably from 1 to 10, and most preferably at least 1 to 5 compared to any one of said amino acid sequences.

15. The protein, composition or method according to any one of claims 1 to 14, having cytotoxic activity against *Klebsiella pneumoniae*, *Klebsiella oxytoca*, *Klebsiella granulomatis*, *Klebsiella quasipneumoniae*, *Klebsiella aerogenes*, and/or *Klebsiella variicola*.

16. The protein, composition or method according to any one of claims 1 to 15, wherein said protein comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 22 to 24, wherein each X stands for any one of the 20 standard amino acid residues or for an absent amino acid residue and each J stands for either L (leucine) or I (isoleucine); or

wherein said protein is any one defined with respect to SEQ ID NOs: 1-4, 7, 8 or 22, preferable any one defined with respect to SEQ ID NOs: 1-4 or 22.

17. The protein, composition or method according to any one of claim 1 to 16, wherein the cytotoxic activity of said protein is such that

- said protein and
- a comparative protein of the amino acid sequence of the SEQ ID NO: 1

produce spots free of viable bacteria of sensitive *Klebsiella quasipneumoniae* subsp. *similipneumoniae* SB30 (DSM 28212) at least of the same diameter 16 hours after spotting 20 microliters of each of a solution of said protein and of the comparative protein onto a lawn of the sensitive *Klebsiella* strain on an agar plate and subsequent incubation of the agar plate at 37°C,

wherein the concentration of said protein in the solution is at most 5 times that of the solution of the comparative protein.

18. A protein having cytotoxic activity against *Klebsiella*, said protein having a lipid II-cleaving activity or a pore-forming capability in a cell membrane of *Klebsiella* cells, said protein preferably comprising or consisting of a first amino acid sequence segment and a second amino acid sequence segment.

19. The protein according to claim 18, further defined as in any one of claims 7 to 17.

20. A composition comprising one or more proteins as defined in any one of claims 7 to 12, 13 to 17, and 18 to 19 and a carrier.

21. A pharmaceutical composition comprising one or more proteins as defined in any one of claims 7 to 12, 13 to 17, and 18 to 19.

22. A composition, preferably a pharmaceutical composition, comprising one or more proteins, wherein said one or more proteins comprise(s) an amino acid sequence selected from the group consisting of SEQ ID NO: 22 to 24, wherein each X stands for any one of the 20 standard amino acid residues or for an absent amino acid residue and each J stands for either L (leucine) or I (isoleucine); or

wherein said protein is any one defined with respect to SEQ ID NOs: 1-4, 7, 8 or 22, preferable any one defined with respect to SEQ ID NOs: 1-4 or 22;

said composition preferably comprising a carrier.

23. The composition according to any one of claims 20 to 22, wherein said composition is a plant material or extract thereof, wherein the plant material may be a material from a plant having expressed said one or more proteins, preferably an edible plant having expressed said one or more proteins.

24. The composition according to any one of claims 20 to 23, wherein said one or more proteins is/are formulated for oral delivery to the small or large intestine.

25. Oral formulation comprising the protein as defined in any one of claims 1 to 17, 18 and 19, or comprising the composition according to any one of claims 20 to 24, said formulation being capable of protecting the protein from gastric conditions and capable of releasing the protein in the intestine.

26. A nucleic acid molecule or nucleic acid construct encoding the protein as defined in any one of claims 1 to 17, 18 or 19, preferably the protein as defined in any one of claims 7 to 17, said nucleic acid molecule or nucleic acid construct comprising a transcription promoter that is preferably active in plant cells and a nucleotide sequence encoding said protein for expressing said nucleotide sequence in cells, preferably in plant cells, under the control of said promoter.

27. A nucleic acid molecule or nucleic acid construct encoding the protein as defined in any one of claims 1 to 17, 18 or 19, preferably the protein as defined in any one of claims 7 to 17, said nucleic acid molecule or nucleic acid construct is or encodes a viral (DNA or RNA)

replicon comprising a nucleotide sequence encoding said protein for expressing said nucleotide sequence in cells, preferably in plant cells.

28. A plant, plant tissue, or plant cell, comprising a protein as defined in any one of claims 1 to 17, 18 and 19, or comprising a nucleic acid molecule or nucleic acid construct as defined in claim 26 or 27.

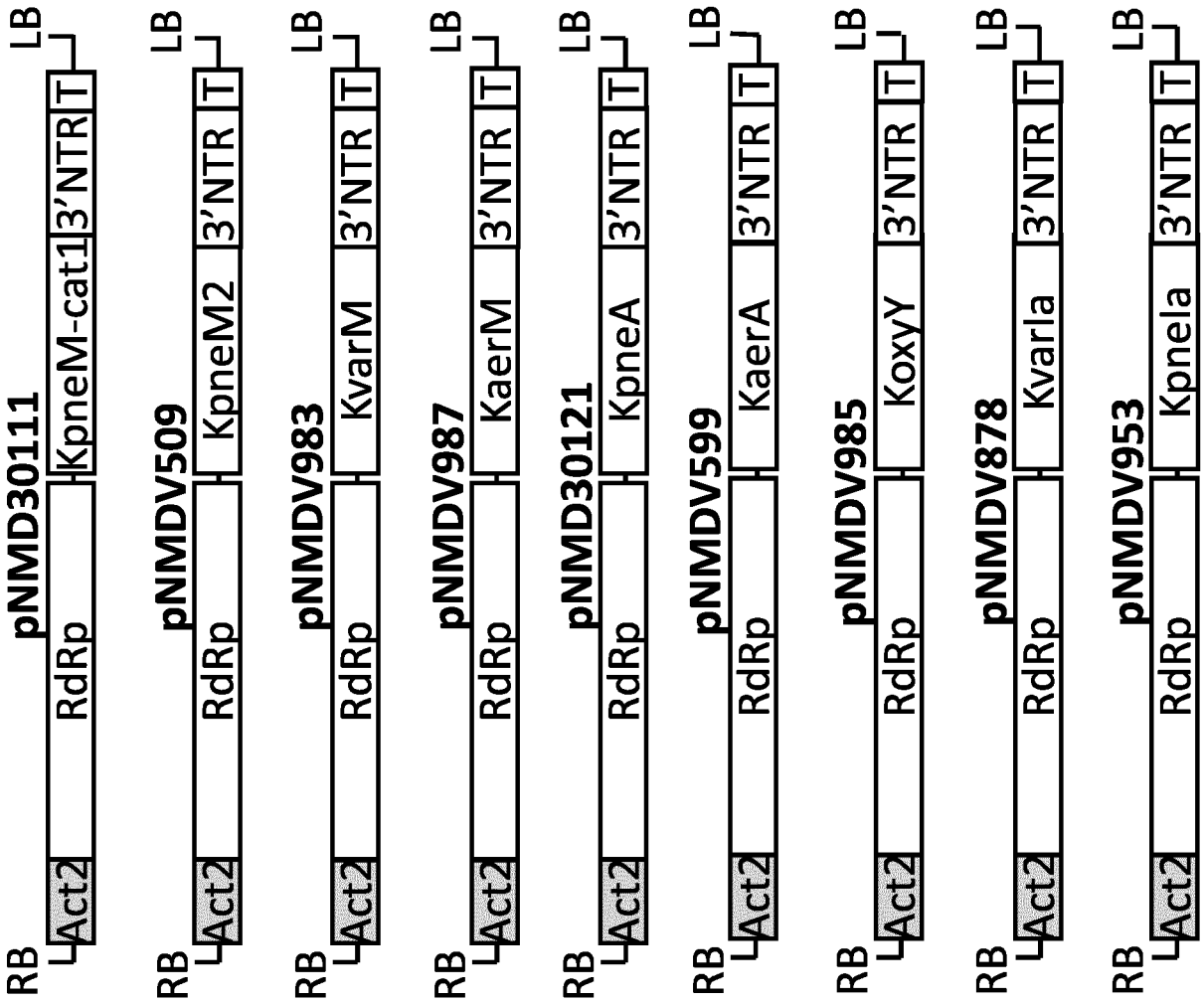

pNMD30111
RB
Act2
RdRp
KpneM-cat1
3'NTR
T
LB
pNMDV509
KpneM2
pNMDV983
KvarM
pNMDV987
KaerM
pNMD30121
KpneA
pNMDV599
KaerA
pNMDV985
KoxyY
pNMDV878
Kvarla
pNMDV953
Kpnela

Fig. 1

Fig. 2

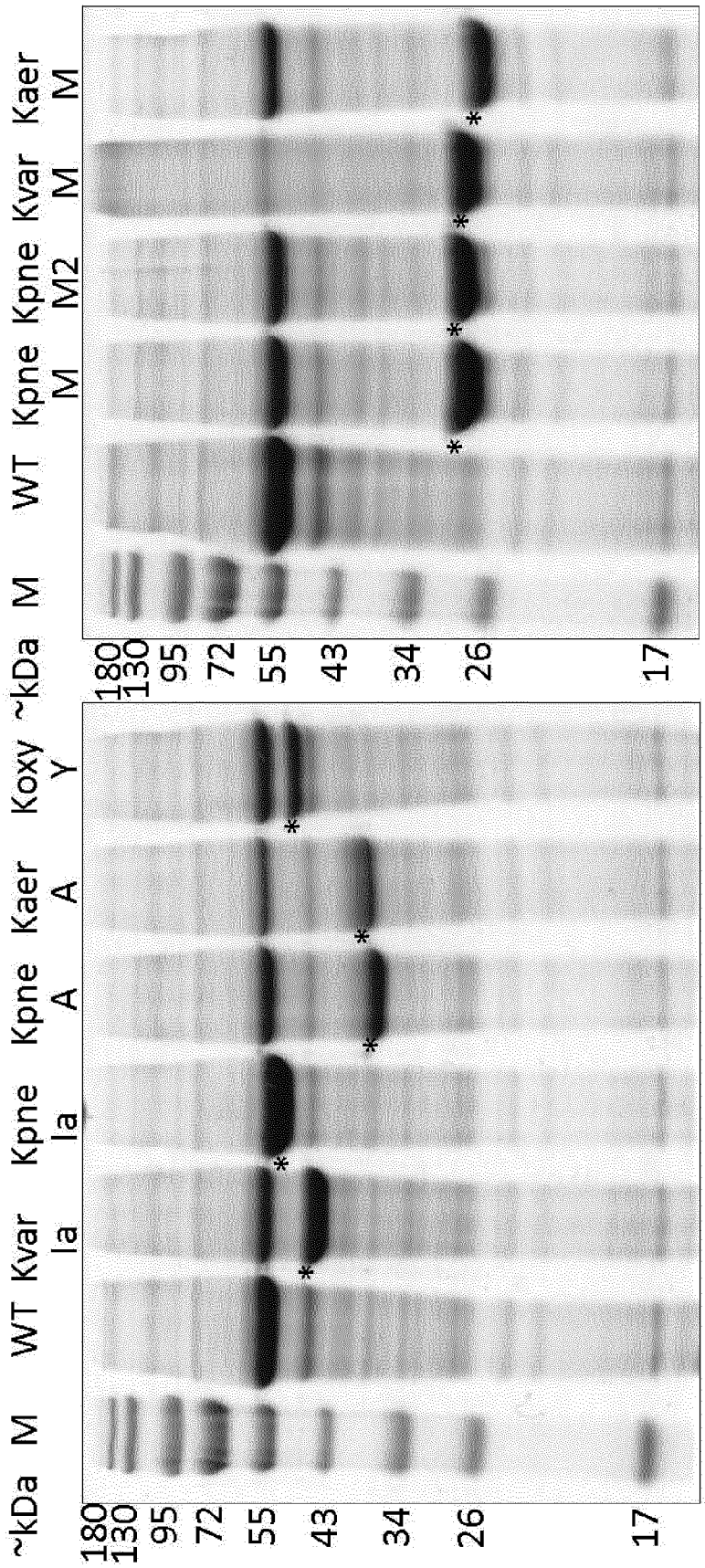
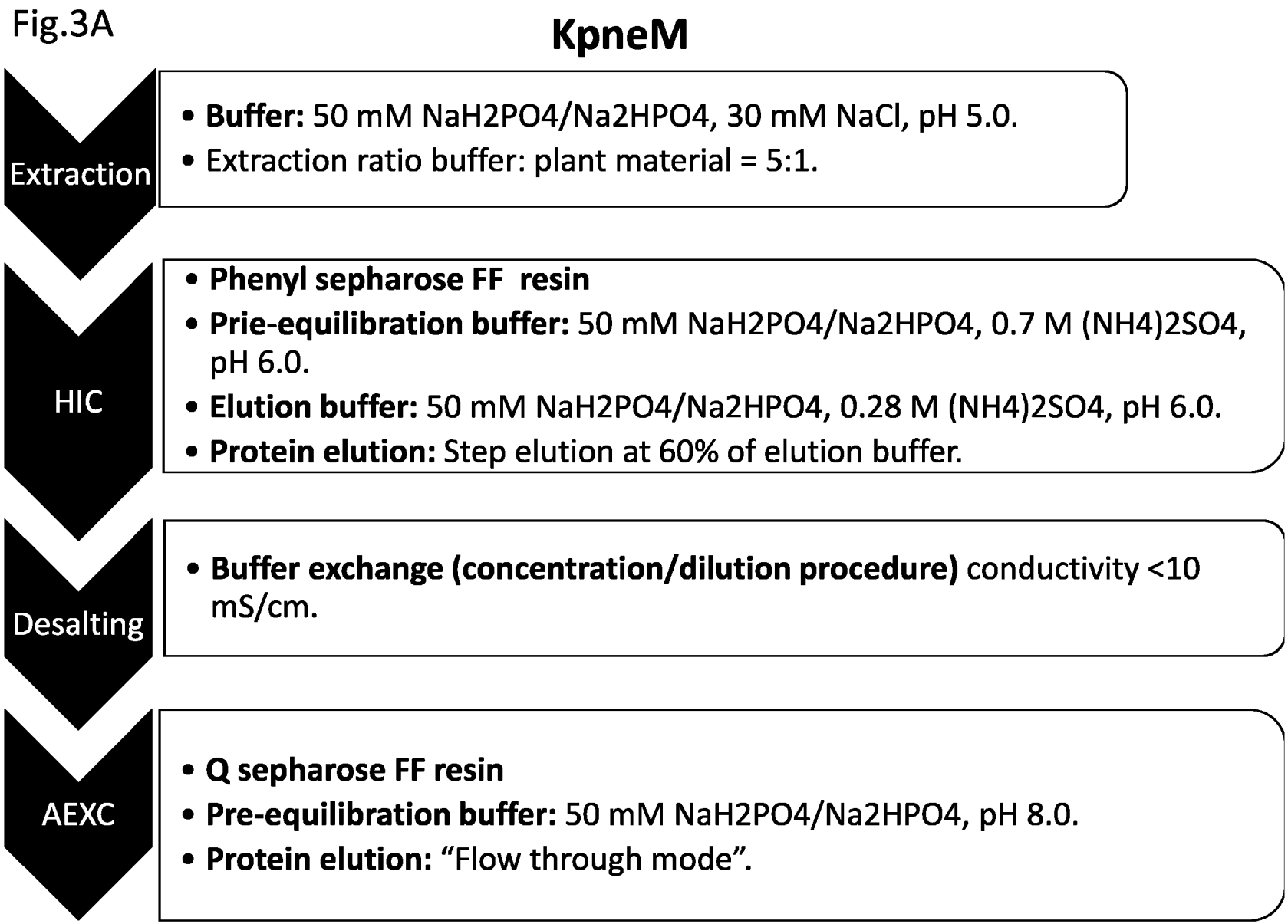

~kDa
M
WT
Kvar 1a
Kpne 1a
Kpne A
Kaer A
Koxy Y
~kDa
M
WT
Kpne M
Kpne M2
Kvar M
Kaer M
180
130
95
72
55
43
34
26
17

Fig.3A


KpneM
Extraction
• Buffer: 50 mM NaH2PO4/Na2HPO4, 30 mM NaCl, pH 5.0.
• Extraction ratio buffer: plant material = 5:1.
HIC
• Phenyl sepharose FF resin
• Prie-equilibration buffer: 50 mM NaH2PO4/Na2HPO4, 0.7 M (NH4)2SO4, pH 6.0.
• Elution buffer: 50 mM NaH2PO4/Na2HPO4, 0.28 M (NH4)2SO4, pH 6.0.
• Protein elution: Step elution at 60% of elution buffer.
Desalting
• Buffer exchange (concentration/dilution procedure) conductivity <10 mS/cm.
AEXC
• Q sepharose FF resin
• Pre-equilibration buffer: 50 mM NaH2PO4/Na2HPO4, pH 8.0.
• Protein elution: "Flow through mode".

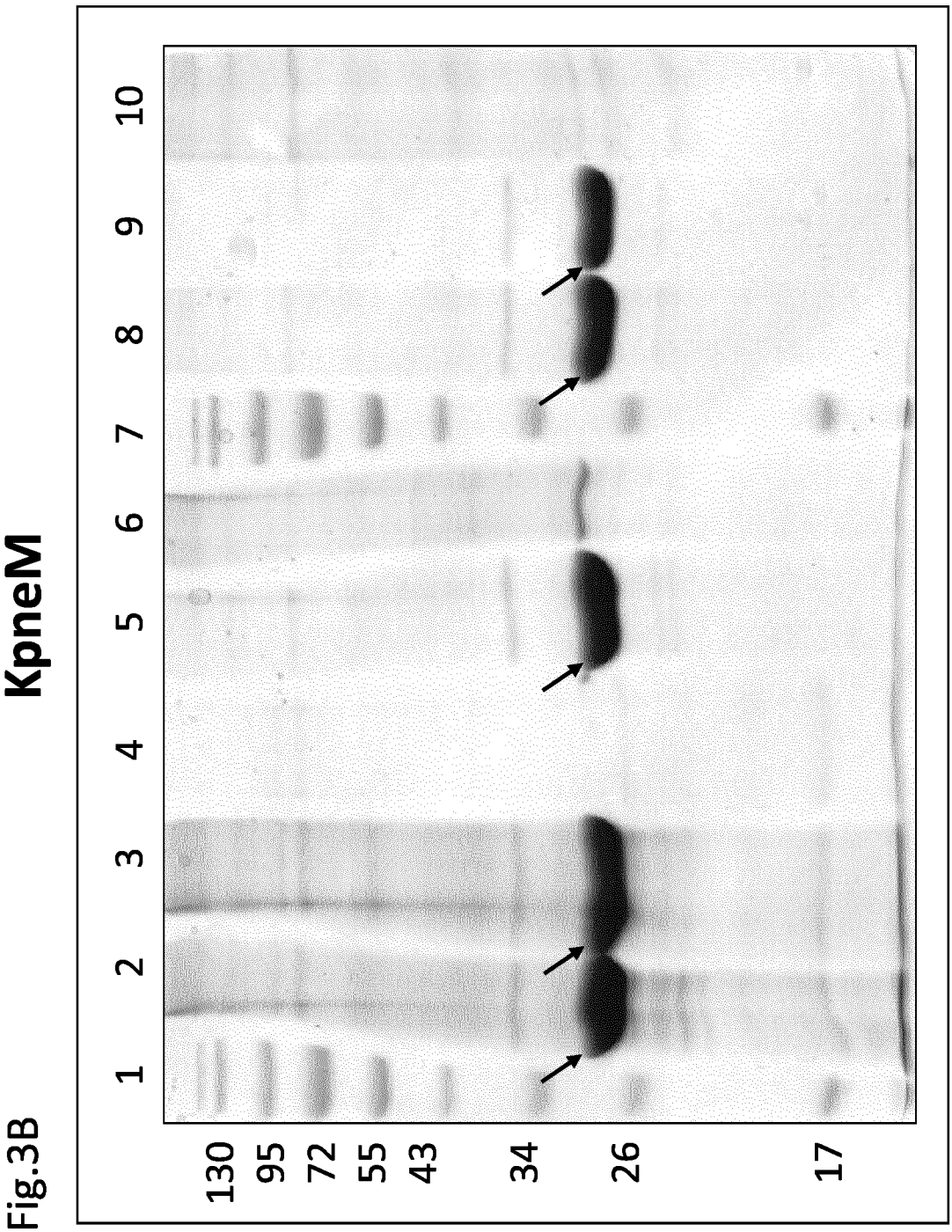



Fig.3B

Fig.3C

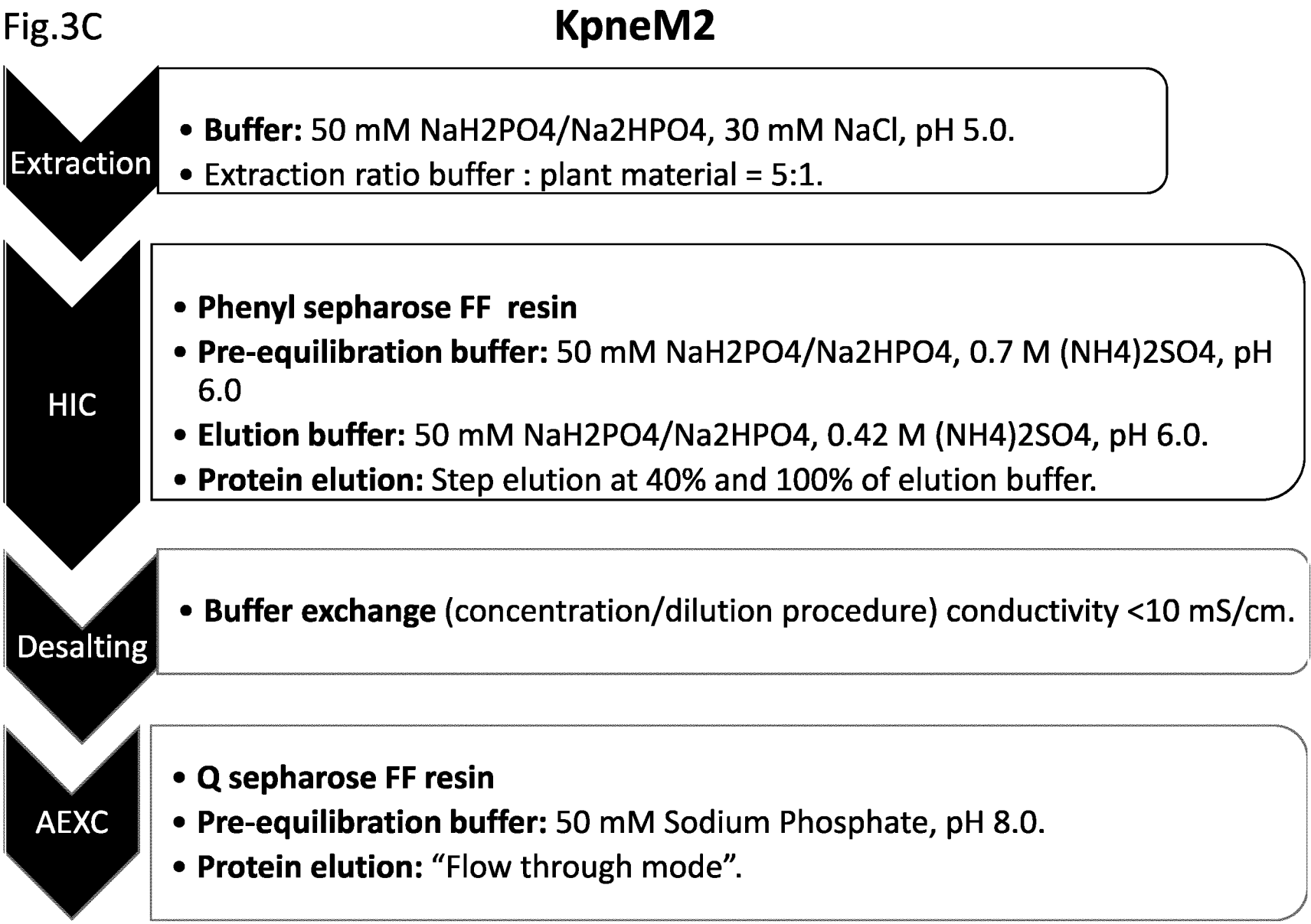

KpneM2
Extraction
• Buffer: 50 mM NaH2PO4/Na2HPO4, 30 mM NaCl, pH 5.0.
• Extraction ratio buffer : plant material = 5:1.
HIC
• Phenyl sepharose FF resin
• Pre-equilibration buffer: 50 mM NaH2PO4/Na2HPO4, 0.7 M (NH4)2SO4, pH 6.0
• Elution buffer: 50 mM NaH2PO4/Na2HPO4, 0.42 M (NH4)2SO4, pH 6.0.
• Protein elution: Step elution at 40% and 100% of elution buffer.
Desalting
• Buffer exchange (concentration/dilution procedure) conductivity <10 mS/cm.
AEXC
• Q sepharose FF resin
• Pre-equilibration buffer: 50 mM Sodium Phosphate, pH 8.0.
• Protein elution: "Flow through mode".

Fig.3D

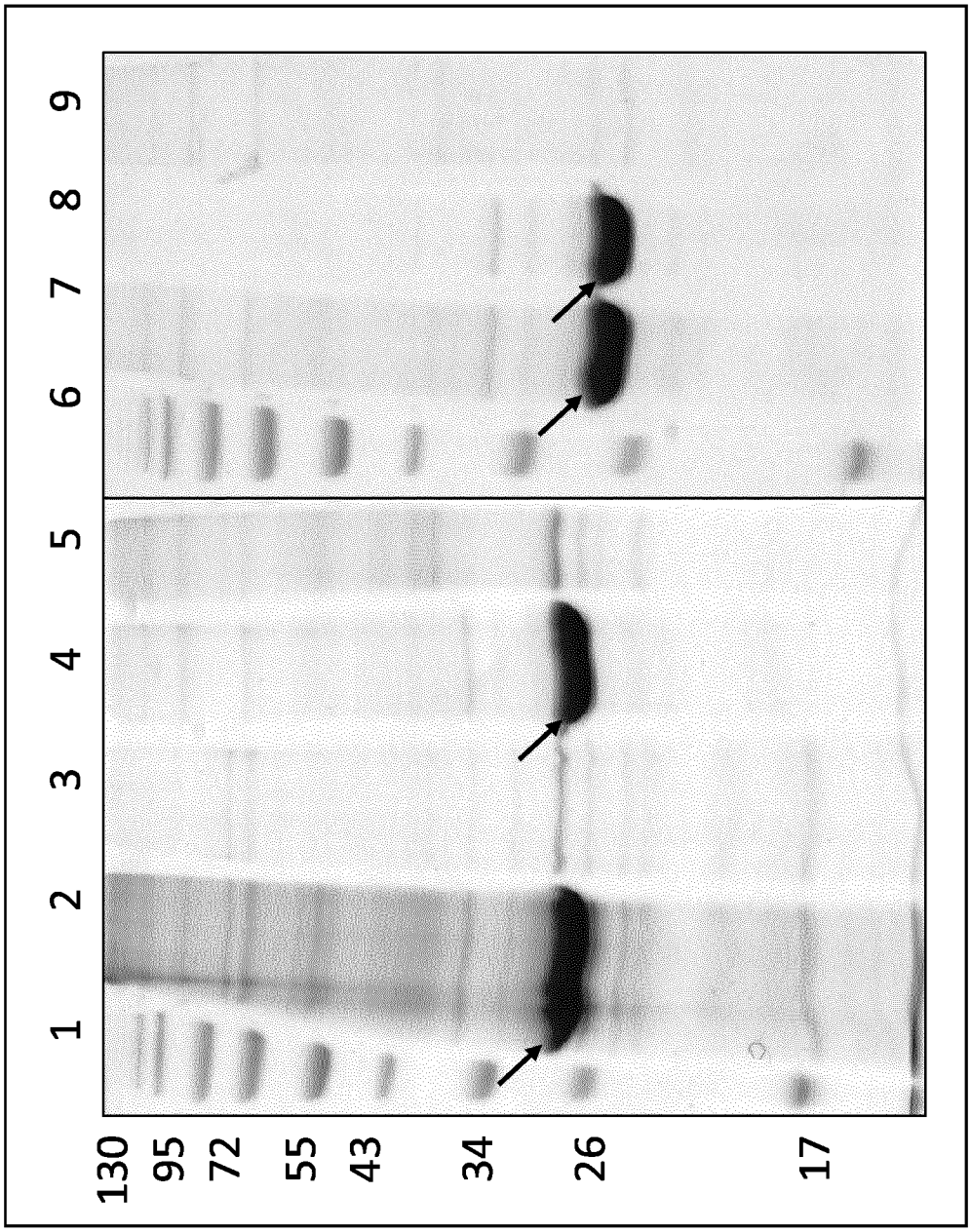
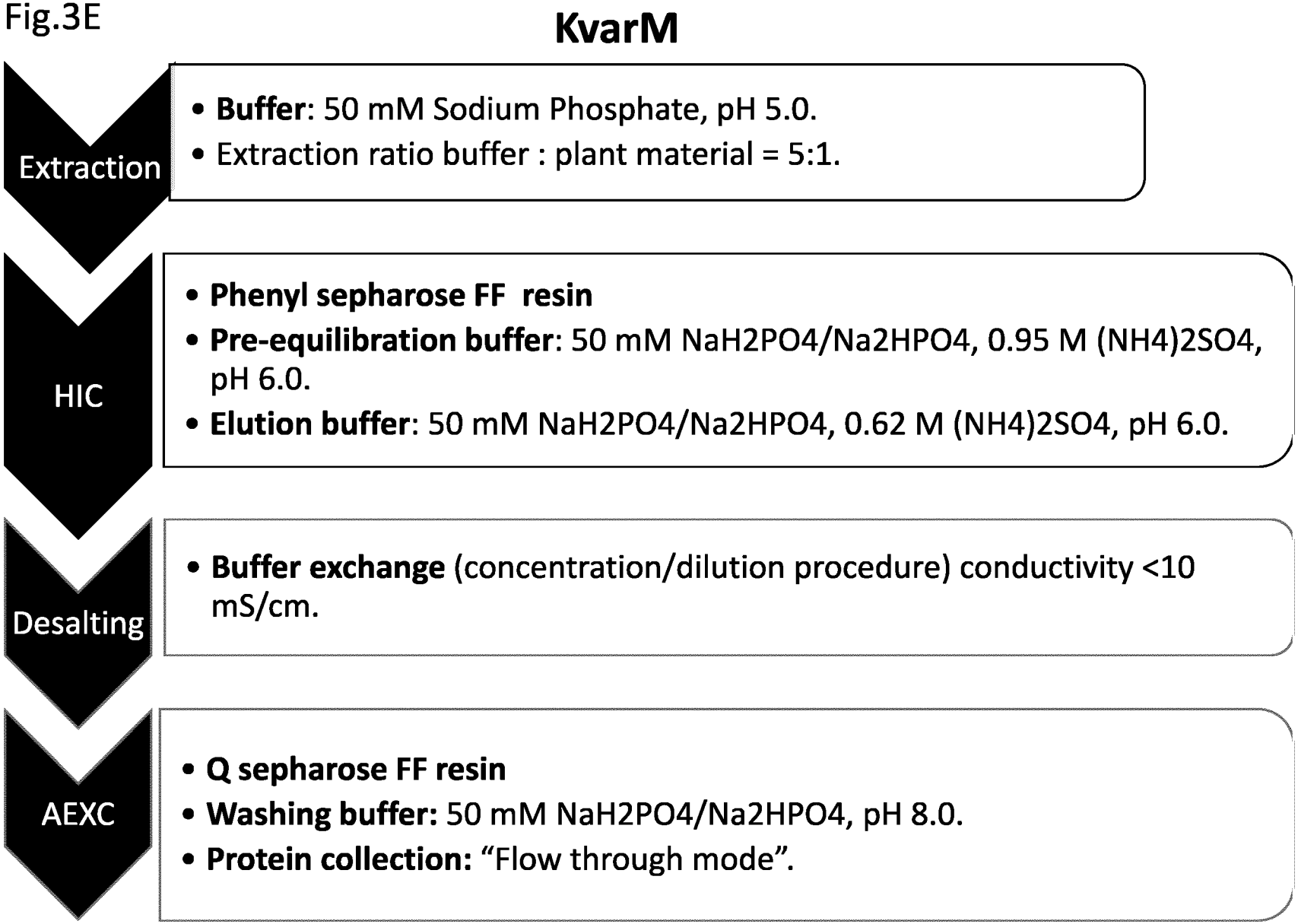


Fig.3E


KvarM
Extraction
• Buffer: 50 mM Sodium Phosphate, pH 5.0.
• Extraction ratio buffer : plant material = 5:1.
HIC
• Phenyl sepharose FF resin
• Pre-equilibration buffer: 50 mM NaH2PO4/Na2HPO4, 0.95 M (NH4)2SO4, pH 6.0.
• Elution buffer: 50 mM NaH2PO4/Na2HPO4, 0.62 M (NH4)2SO4, pH 6.0.
Desalting
• Buffer exchange (concentration/dilution procedure) conductivity <10 mS/cm.
AEXC
• Q sepharose FF resin
• Washing buffer: 50 mM NaH2PO4/Na2HPO4, pH 8.0.
• Protein collection: "Flow through mode".

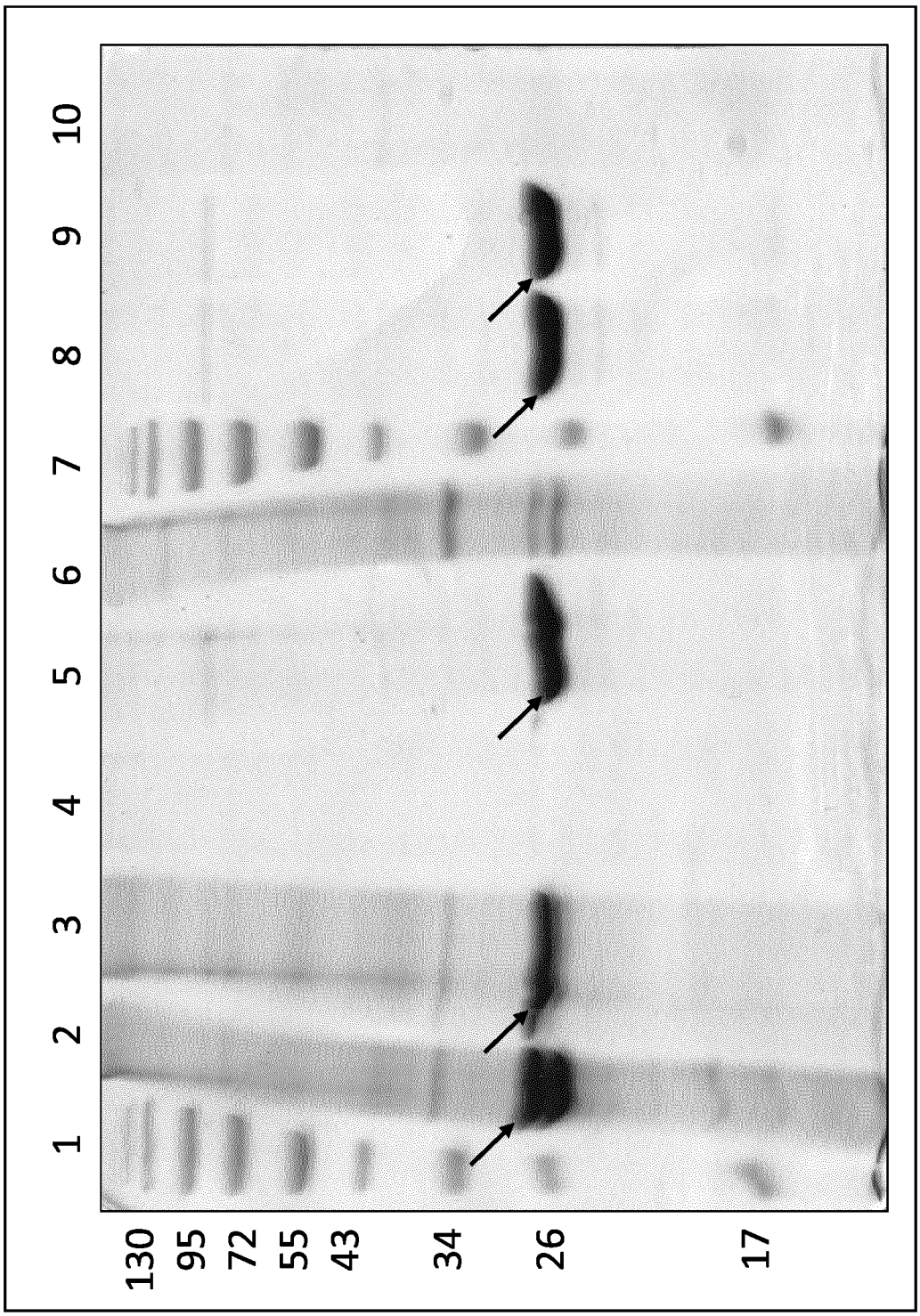
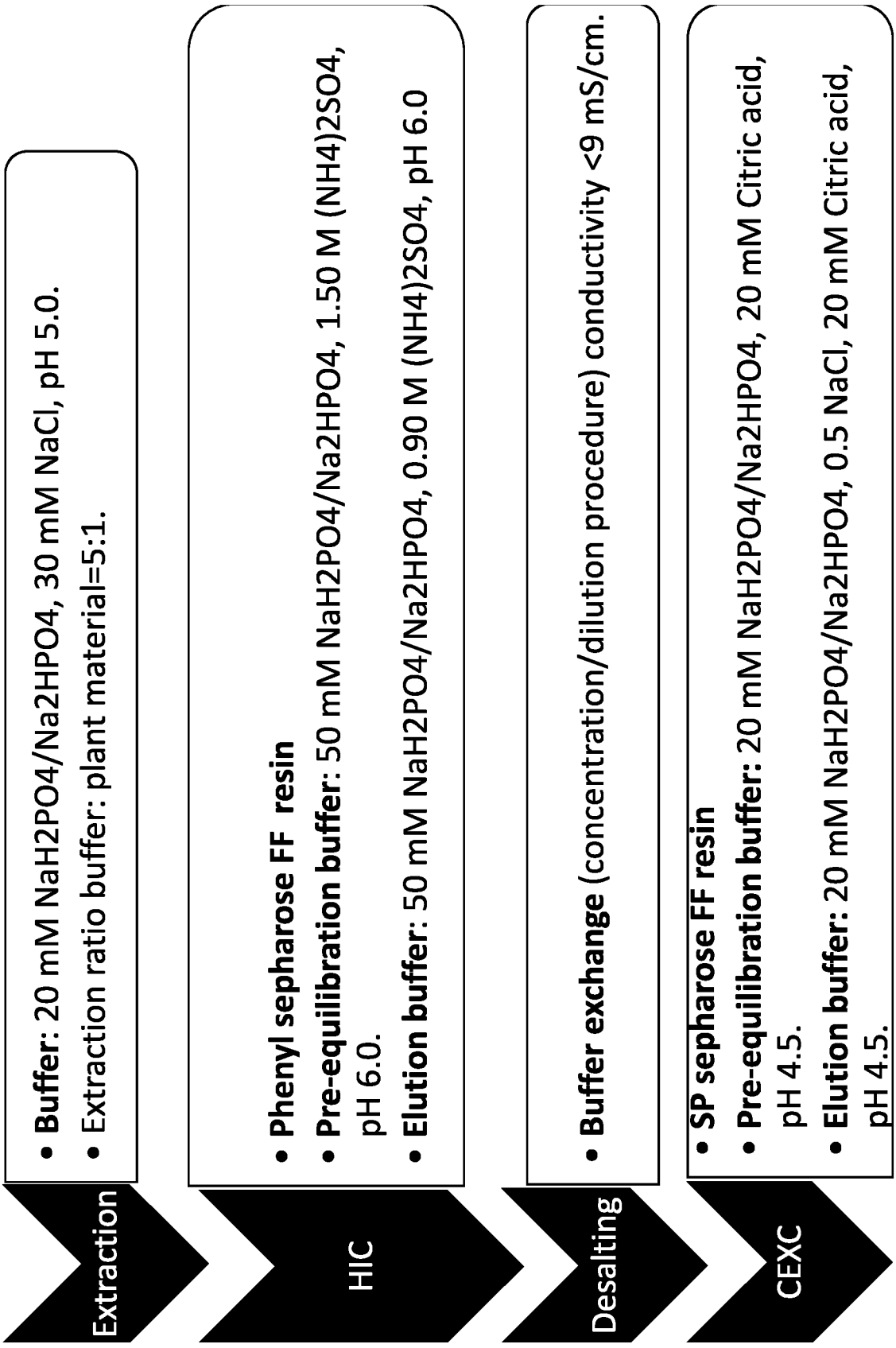



Fig.3F

Fig.3G


KpneA
Extraction
• Buffer: 20 mM NaH2PO4/Na2HPO4, 30 mM NaCl, pH 5.0.
• Extraction ratio buffer: plant material=5:1.
HIC
• Phenyl sepharose FF resin
• Pre-equilibration buffer: 50 mM NaH2PO4/Na2HPO4, 1.50 M (NH4)2SO4, pH 6.0.
• Elution buffer: 50 mM NaH2PO4/Na2HPO4, 0.90 M (NH4)2SO4, pH 6.0
Desalting
• Buffer exchange (concentration/dilution procedure) conductivity <9 mS/cm.
CEXC
• SP sepharose FF resin
• Pre-equilibration buffer: 20 mM NaH2PO4/Na2HPO4, 20 mM Citric acid, pH 4.5.
• Elution buffer: 20 mM NaH2PO4/Na2HPO4, 0.5 NaCl, 20 mM Citric acid, pH 4.5.

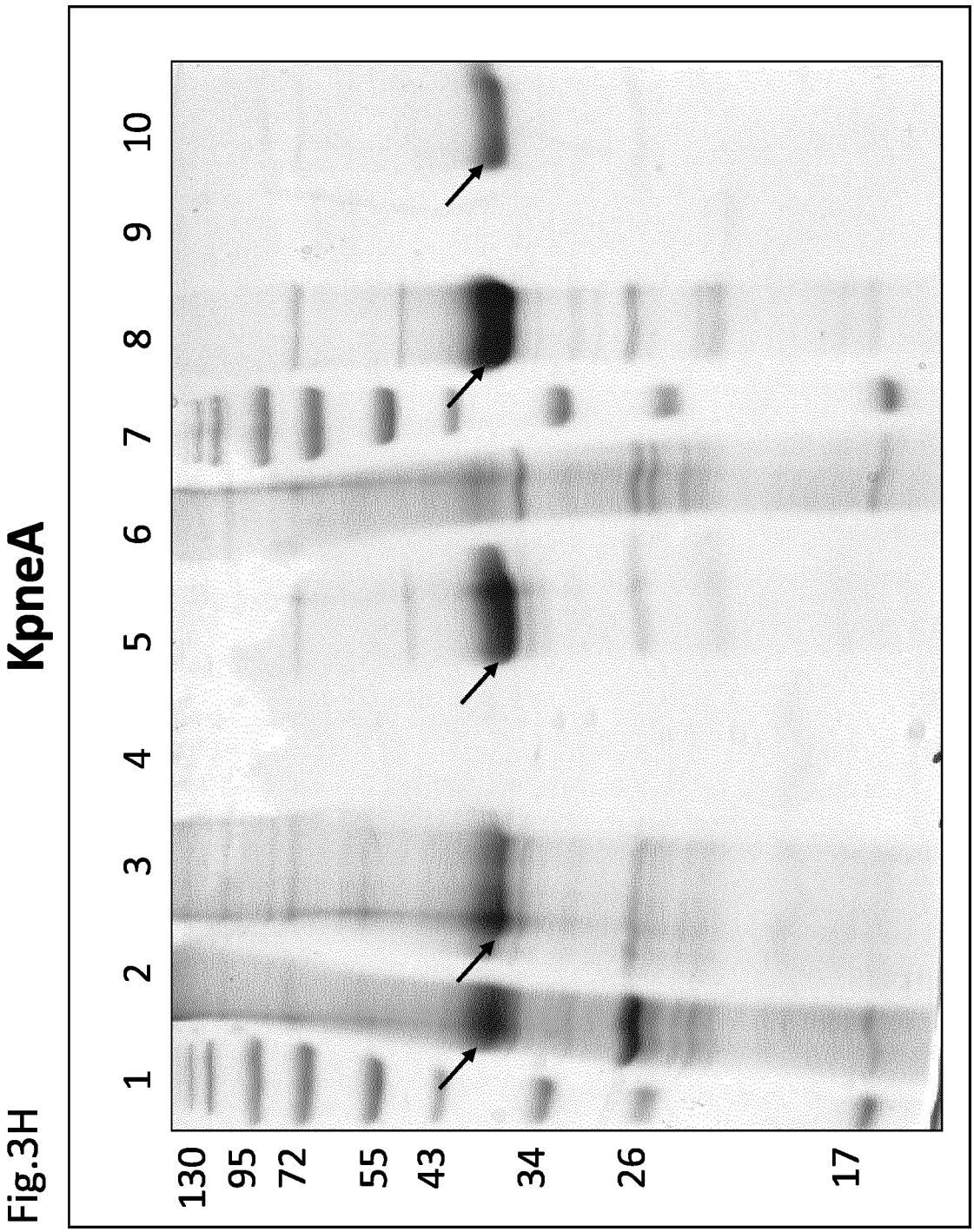
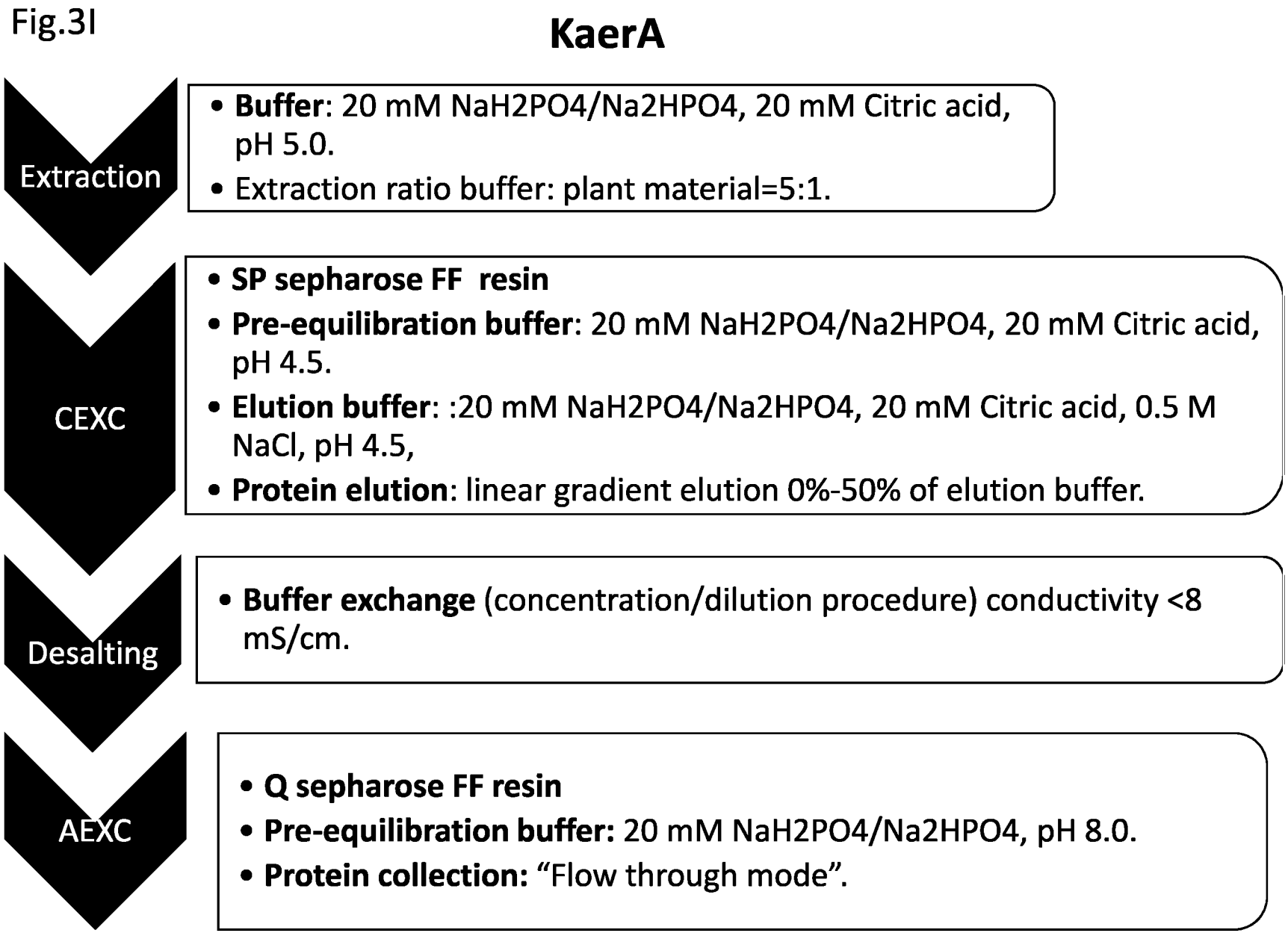



Fig.3H

Fig.3I


KaerA
Extraction
• Buffer: 20 mM NaH2PO4/Na2HPO4, 20 mM Citric acid, pH 5.0.
• Extraction ratio buffer: plant material=5:1.
CEXC
• SP sepharose FF resin
• Pre-equilibration buffer: 20 mM NaH2PO4/Na2HPO4, 20 mM Citric acid, pH 4.5.
• Elution buffer: :20 mM NaH2PO4/Na2HPO4, 20 mM Citric acid, 0.5 M NaCl, pH 4.5,
• Protein elution: linear gradient elution 0%-50% of elution buffer.
Desalting
• Buffer exchange (concentration/dilution procedure) conductivity <8 mS/cm.
AEXC
• Q sepharose FF resin
• Pre-equilibration buffer: 20 mM NaH2PO4/Na2HPO4, pH 8.0.
• Protein collection: "Flow through mode".

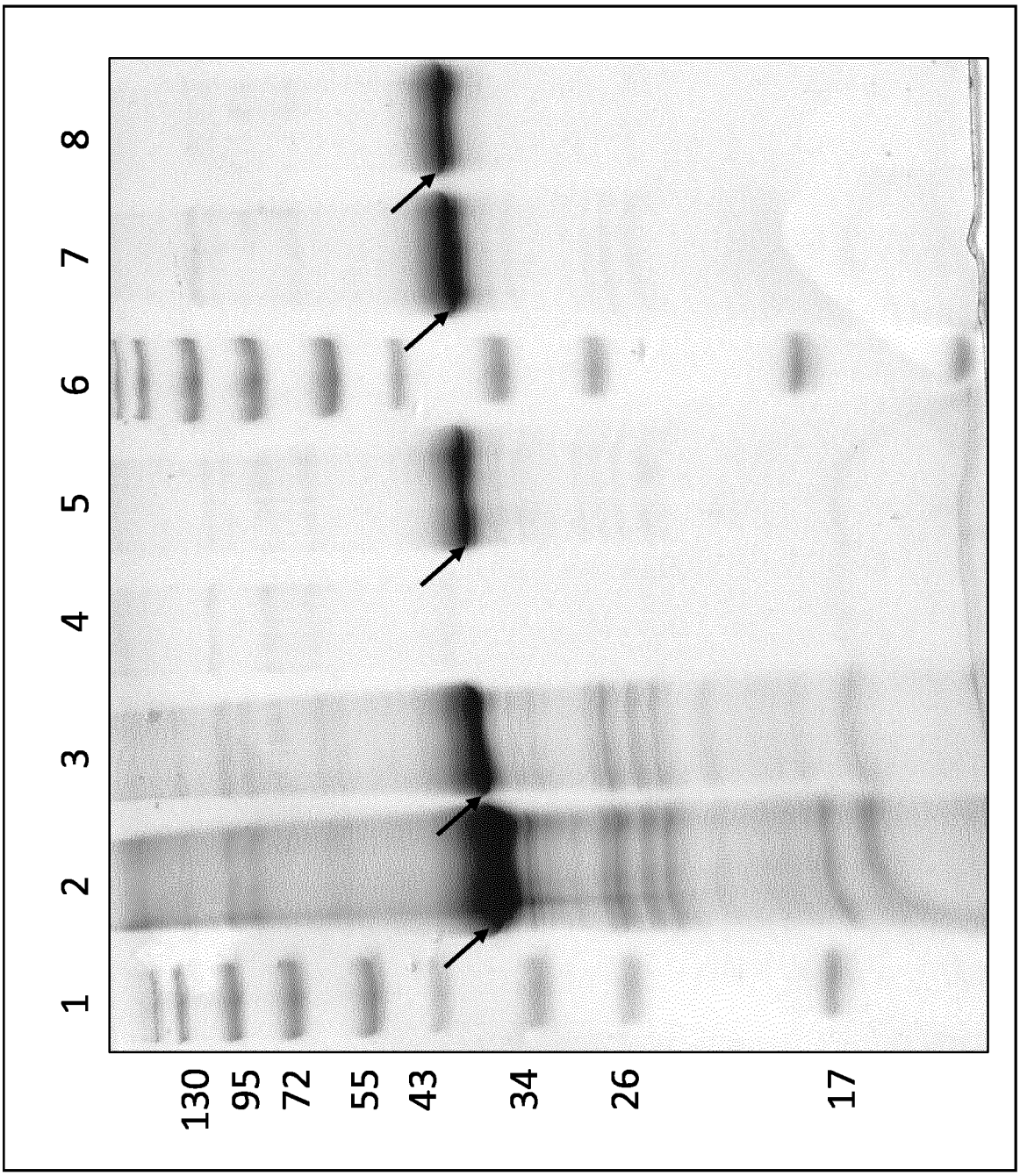
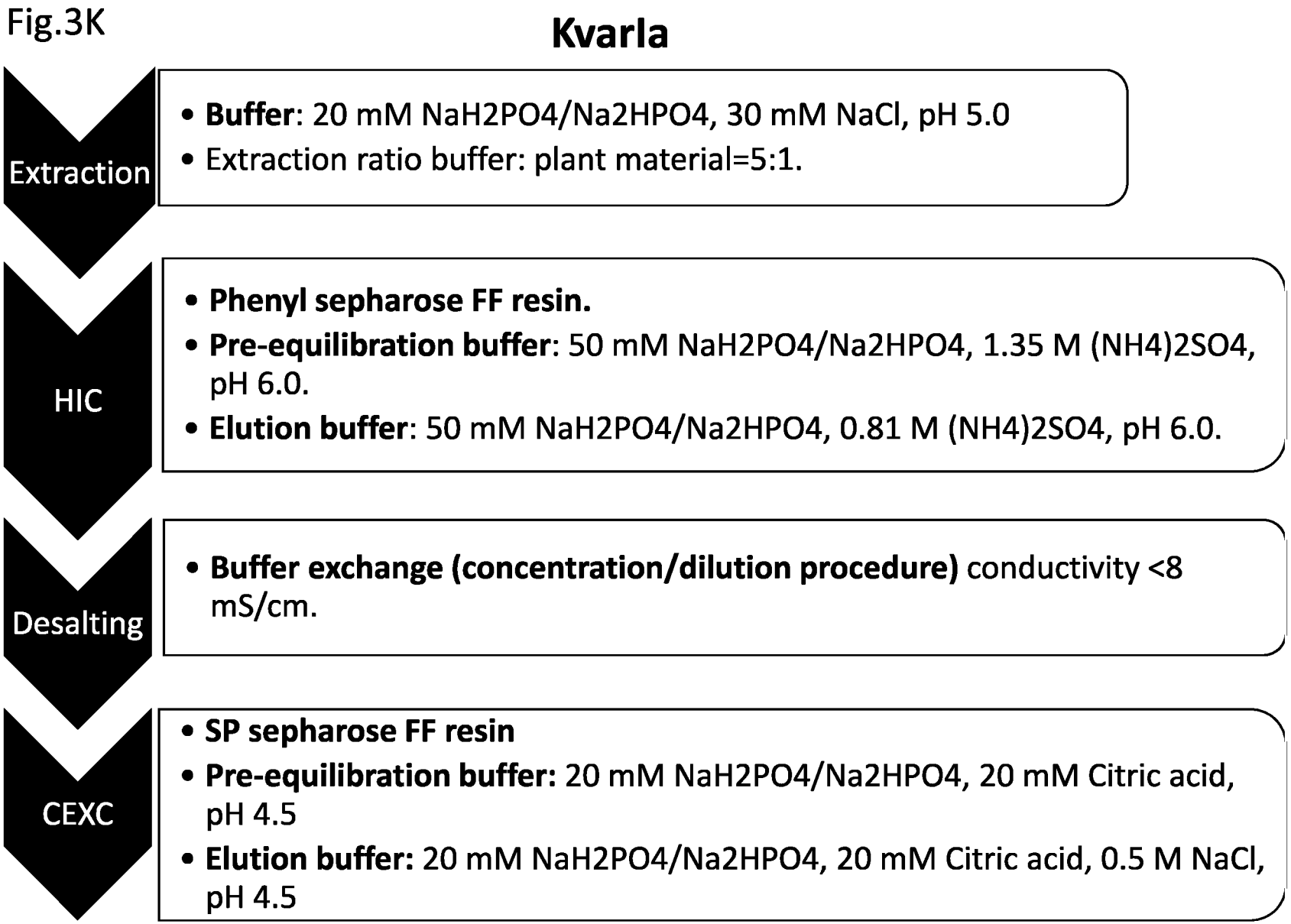



Fig.3J

Fig.3K


Kvarla
Extraction
• **Buffer**: 20 mM NaH2PO4/Na2HPO4, 30 mM NaCl, pH 5.0
• Extraction ratio buffer: plant material=5:1.
HIC
• **Phenyl sepharose FF resin.**
• **Pre-equilibration buffer**: 50 mM NaH2PO4/Na2HPO4, 1.35 M (NH4)2SO4, pH 6.0.
• **Elution buffer**: 50 mM NaH2PO4/Na2HPO4, 0.81 M (NH4)2SO4, pH 6.0.
Desalting
• **Buffer exchange (concentration/dilution procedure)** conductivity <8 mS/cm.
CEXC
• **SP sepharose FF resin**
• **Pre-equilibration buffer:** 20 mM NaH2PO4/Na2HPO4, 20 mM Citric acid, pH 4.5
• **Elution buffer:** 20 mM NaH2PO4/Na2HPO4, 20 mM Citric acid, 0.5 M NaCl, pH 4.5

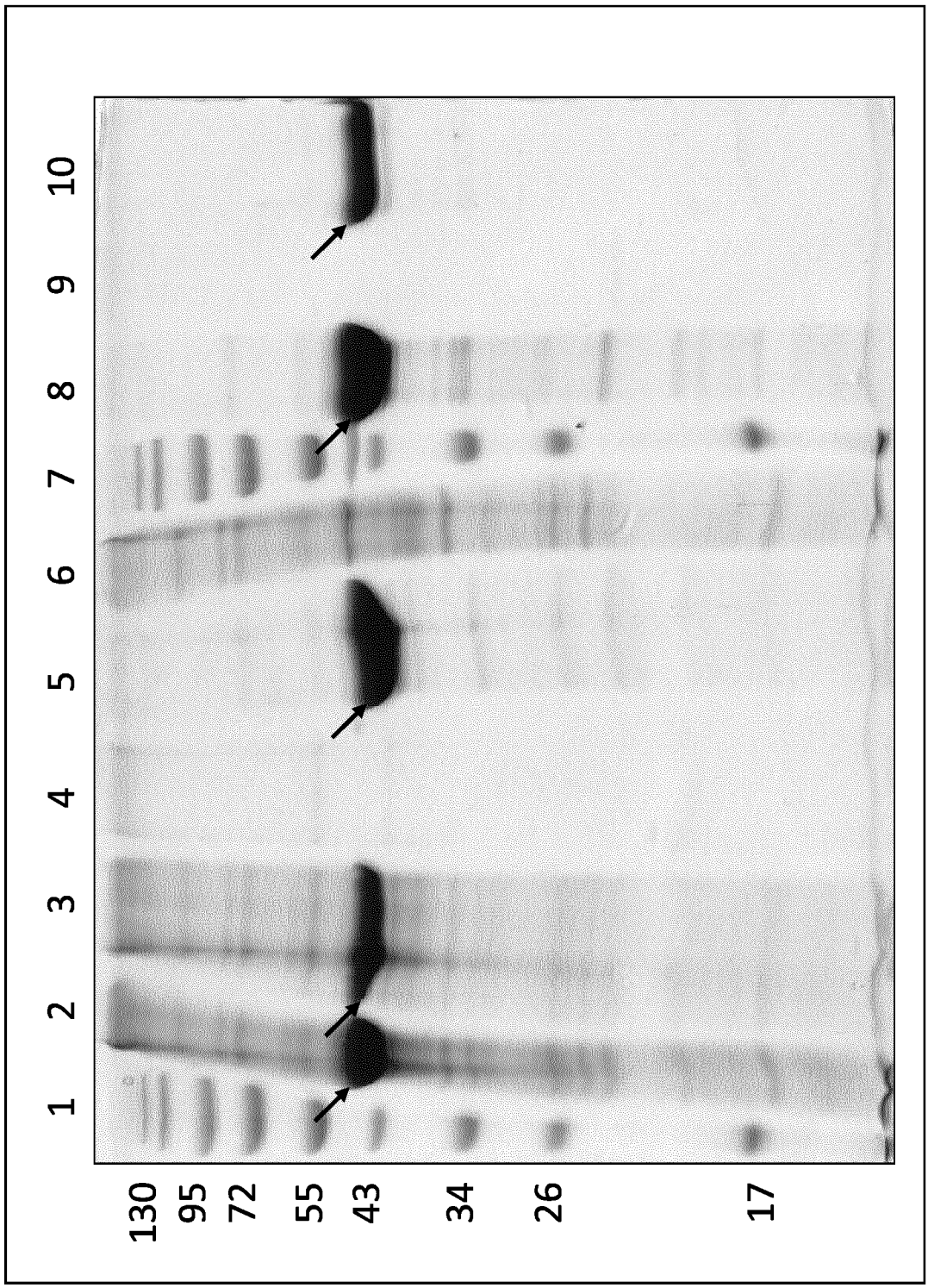



Fig.3L

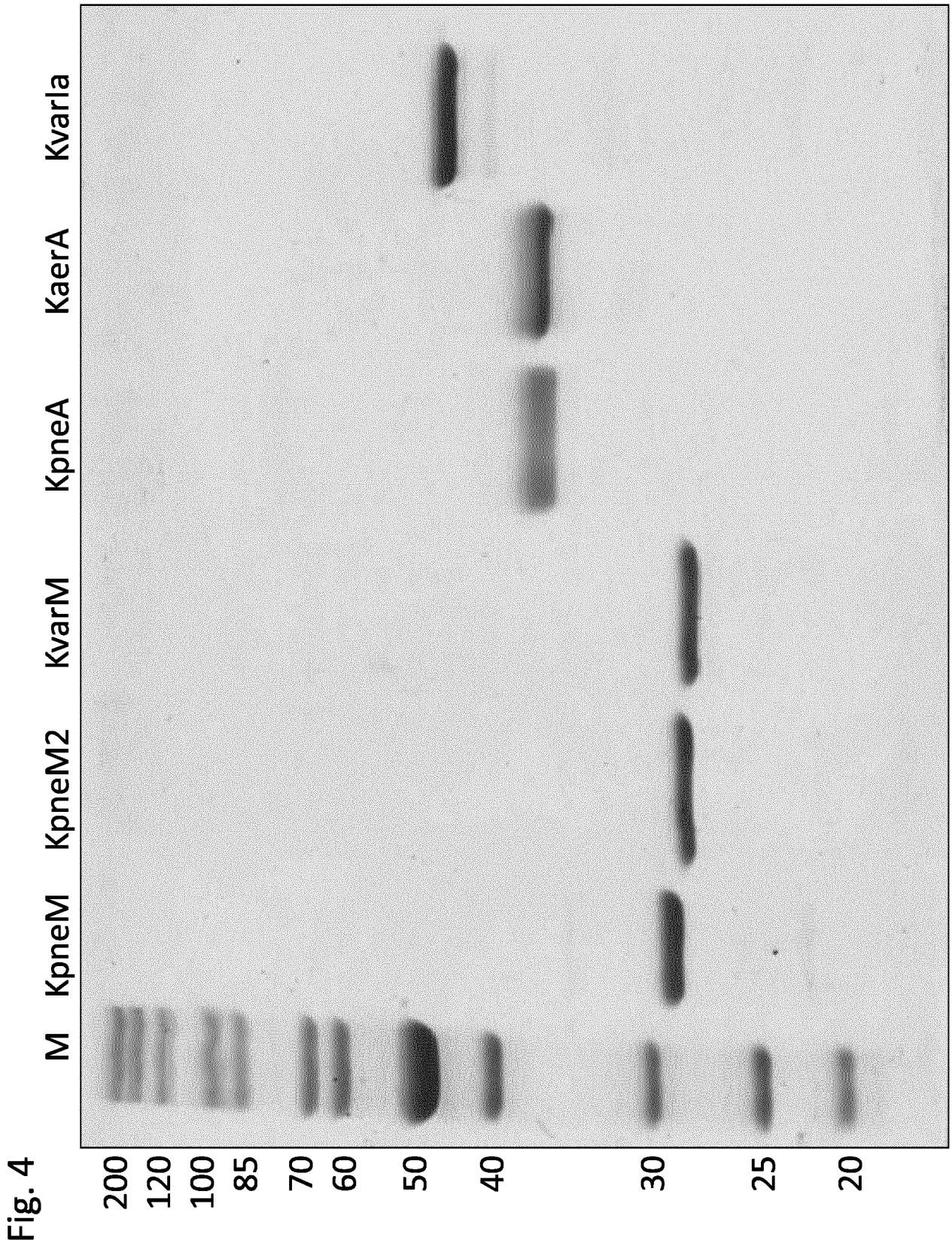

M
KpneM
KpneM2
KvarM
KpneA
KaerA
Kvarla
200
120
100
85
70
60
50
40
30
25
20

Fig. 4

Fig. 5

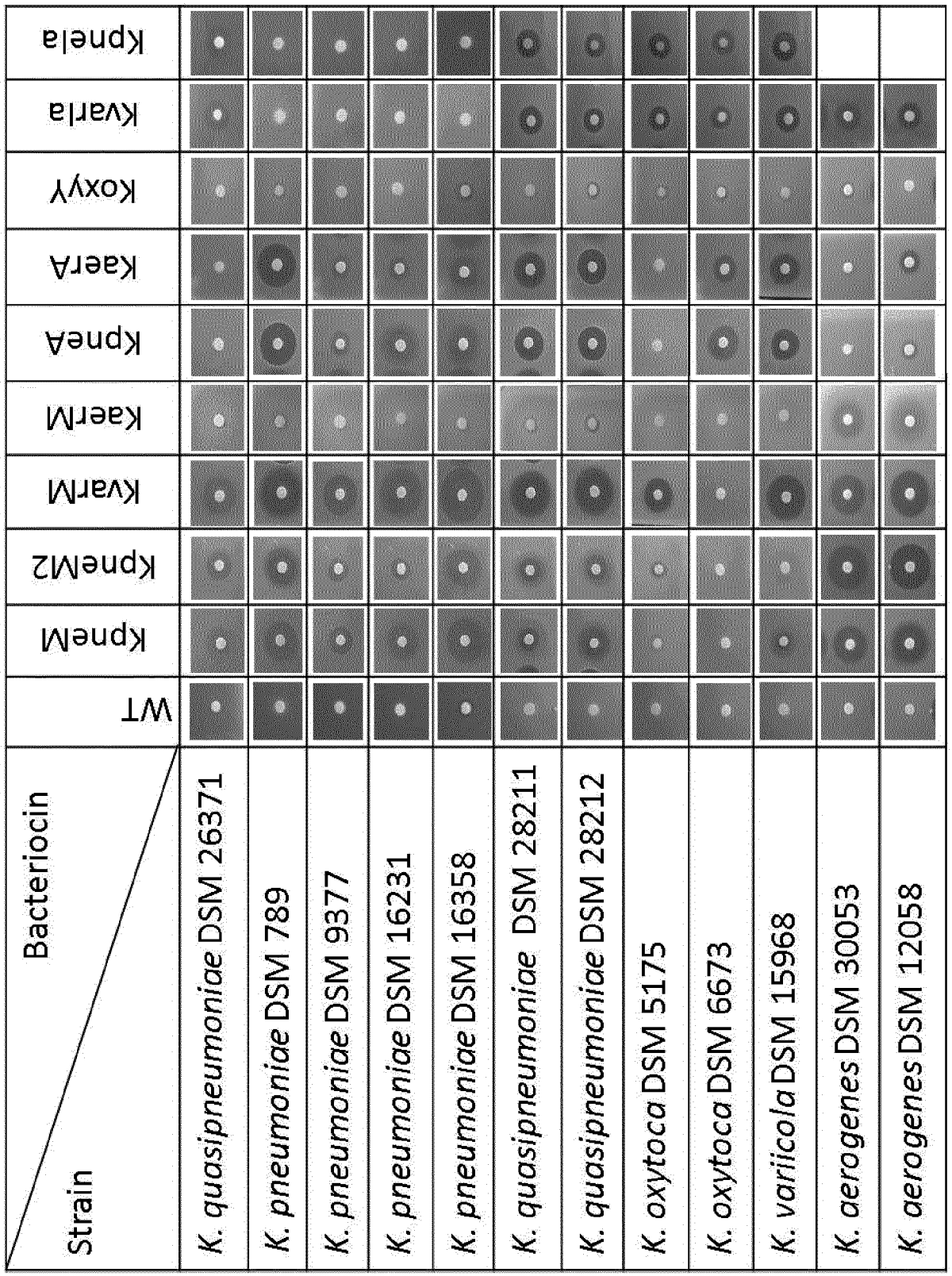
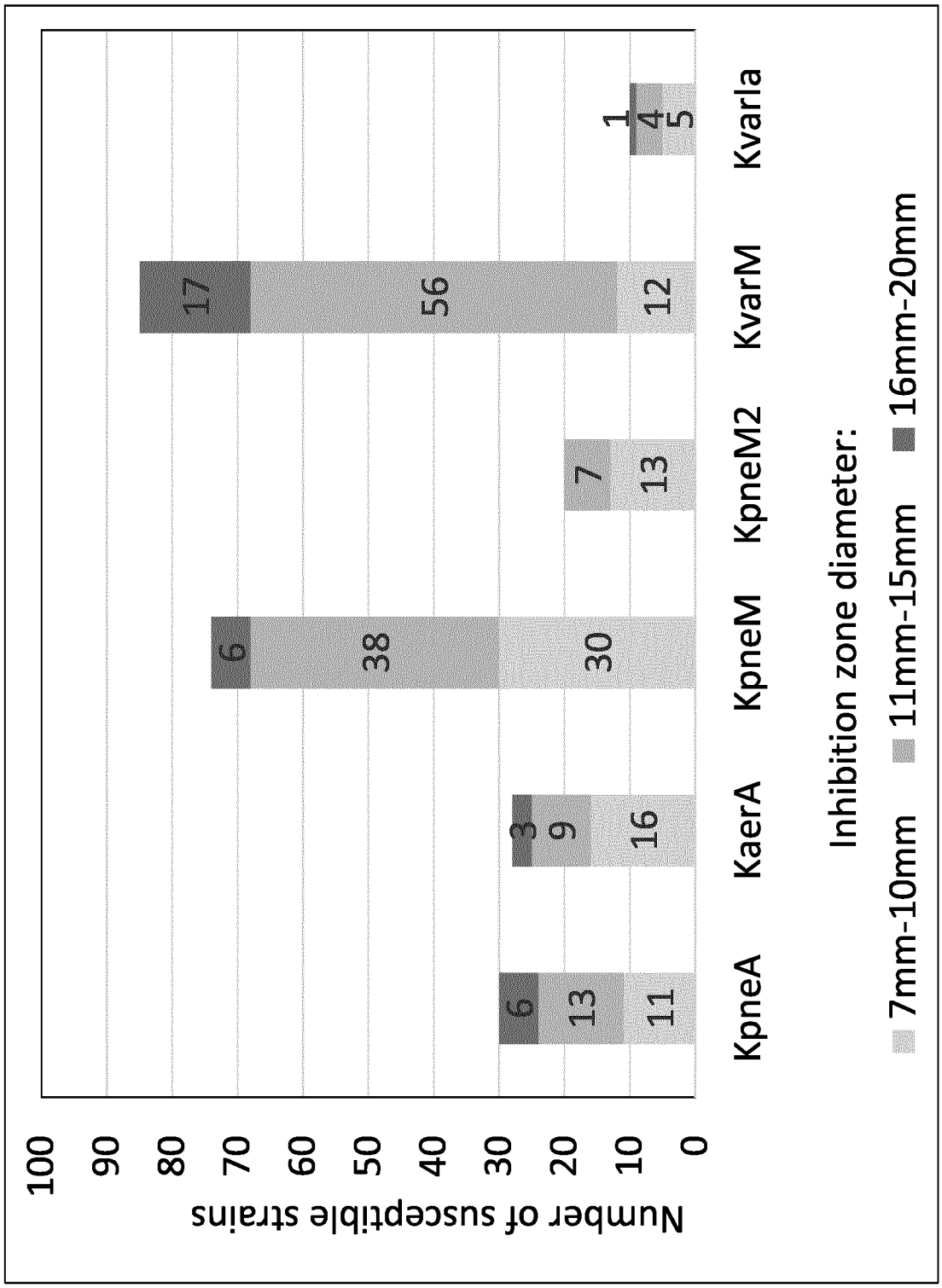

Bacteriocin
Strain
WT
KpneM
KpneM2
KvarM
KaerM
KpneA
KaerA
KoxyY
Kvarla
Kpnela
K. quasipneumoniae DSM 26371
K. pneumoniae DSM 789
K. pneumoniae DSM 9377
K. pneumoniae DSM 16231
K. pneumoniae DSM 16358
K. quasipneumoniae DSM 28211
K. quasipneumoniae DSM 28212
K. oxytoca DSM 5175
K. oxytoca DSM 6673
K. variicola DSM 15968
K. aerogenes DSM 30053
K. aerogenes DSM 12058

Fig. 6


Number of susceptible strains
100
90
80
70
60
50
40
30
20
10
0
KpneA
KaerA
KpneM
KpneM2
KvarM
Kvarla
6
13
11
3
9
16
6
38
30
7
13
17
56
12
1
4
5
Inhibition zone diameter:
7mm-10mm
11mm-15mm
16mm-20mm

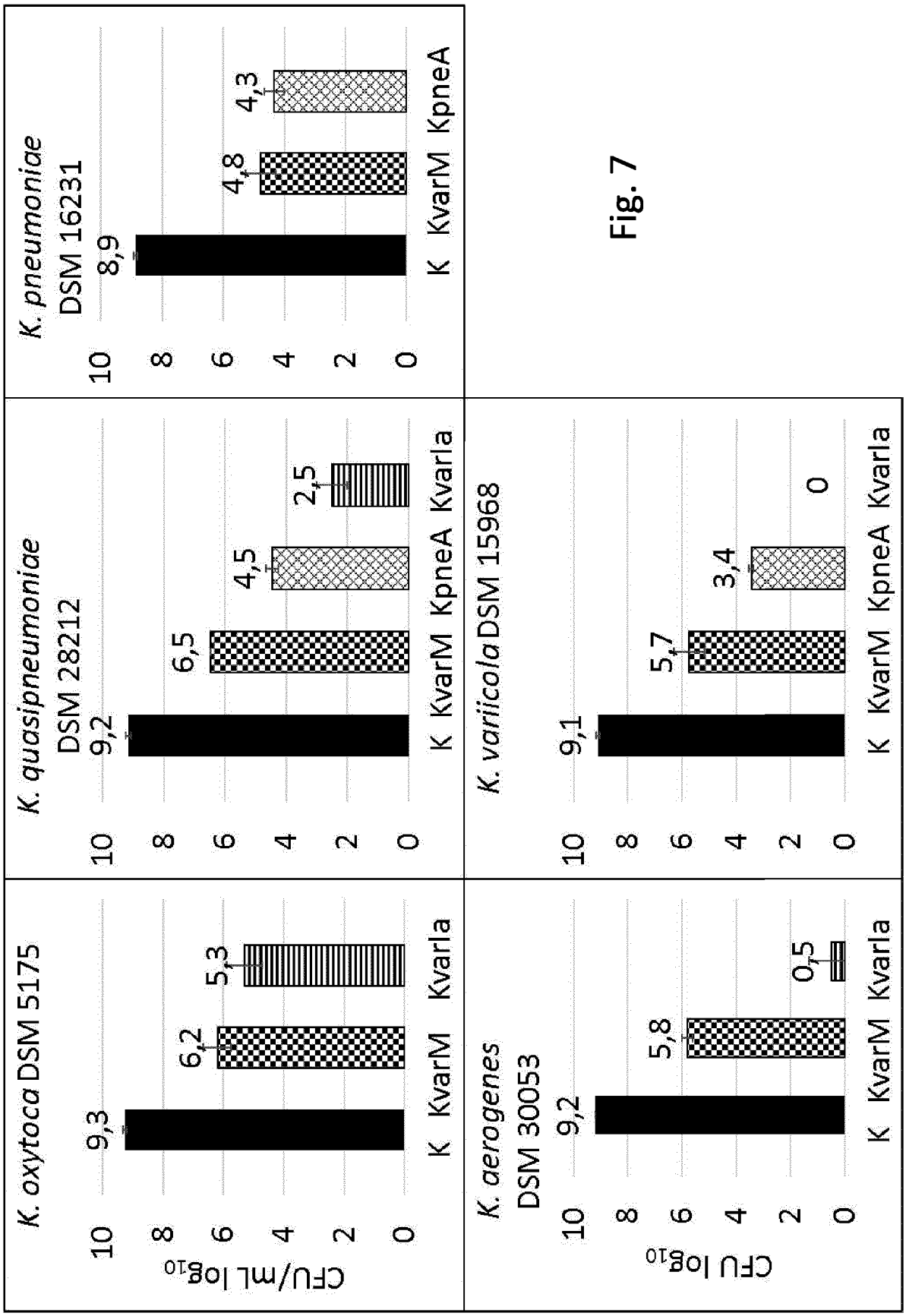

K. oxytoca DSM 5175
K. quasipneumoniae DSM 28212
K. pneumoniae DSM 16231
K. aerogenes DSM 30053
K. variicola DSM 15968
CFU/mL log10
CFU log10
9,3
6,2
5,3
9,2
6,5
4,5
2,5
8,9
4,8
4,3
9,2
5,8
0,5
9,1
5,7
3,4
0
K KvarM KpneA Kvarla

Fig. 7

Fig. 8

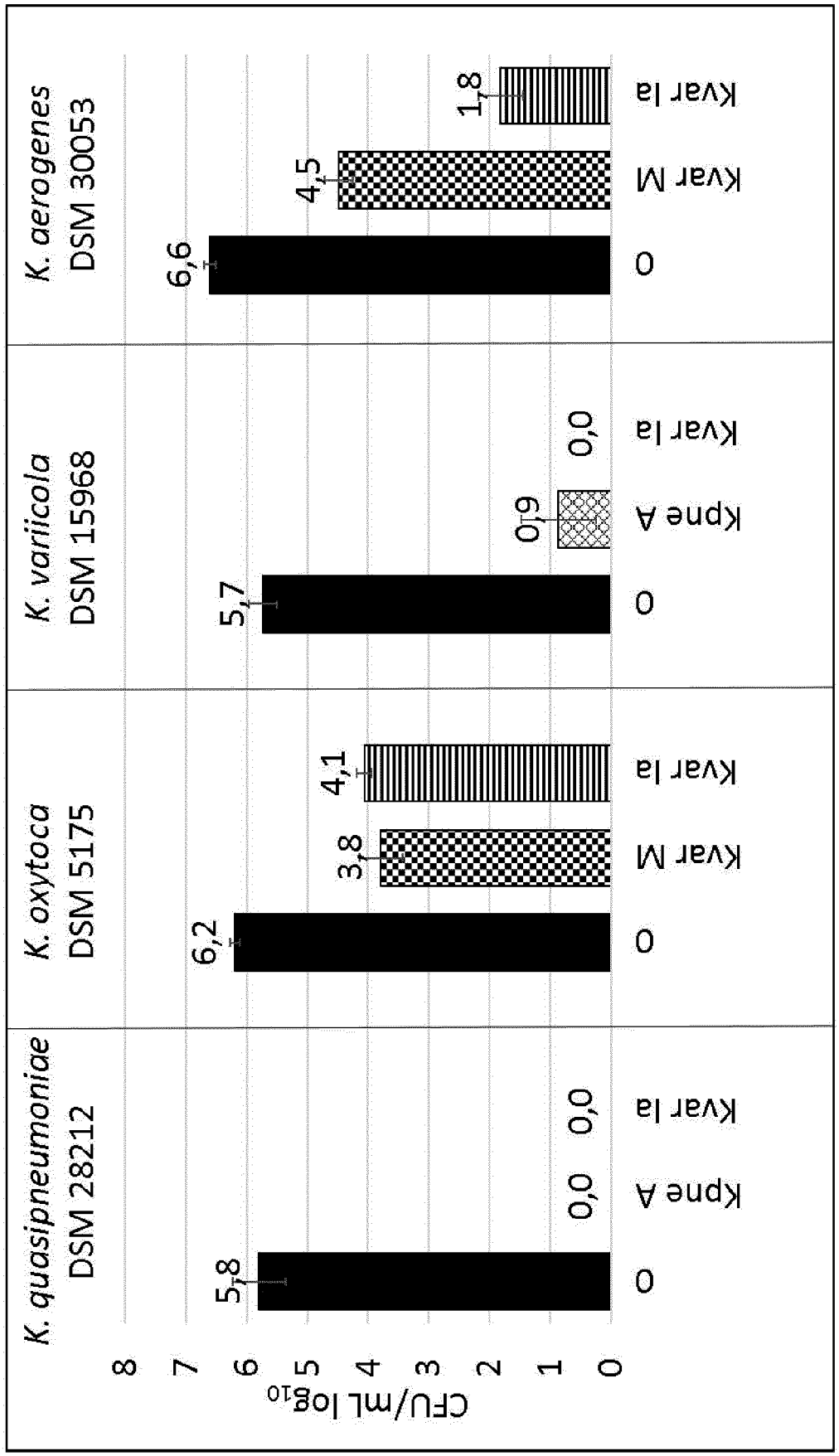

K. quasipneumoniae DSM 28212
K. oxytoca DSM 5175
K. variicola DSM 15968
K. aerogenes DSM 30053
CFU/mL log10
0
1
2
3
4
5
6
7
8
5,8
0,0
0,0
6,2
3,8
4,1
5,7
6,0
0,0
6,6
4,5
1,8
0
Kpne A
Kvar M
Kvar Ia

Fig. 9

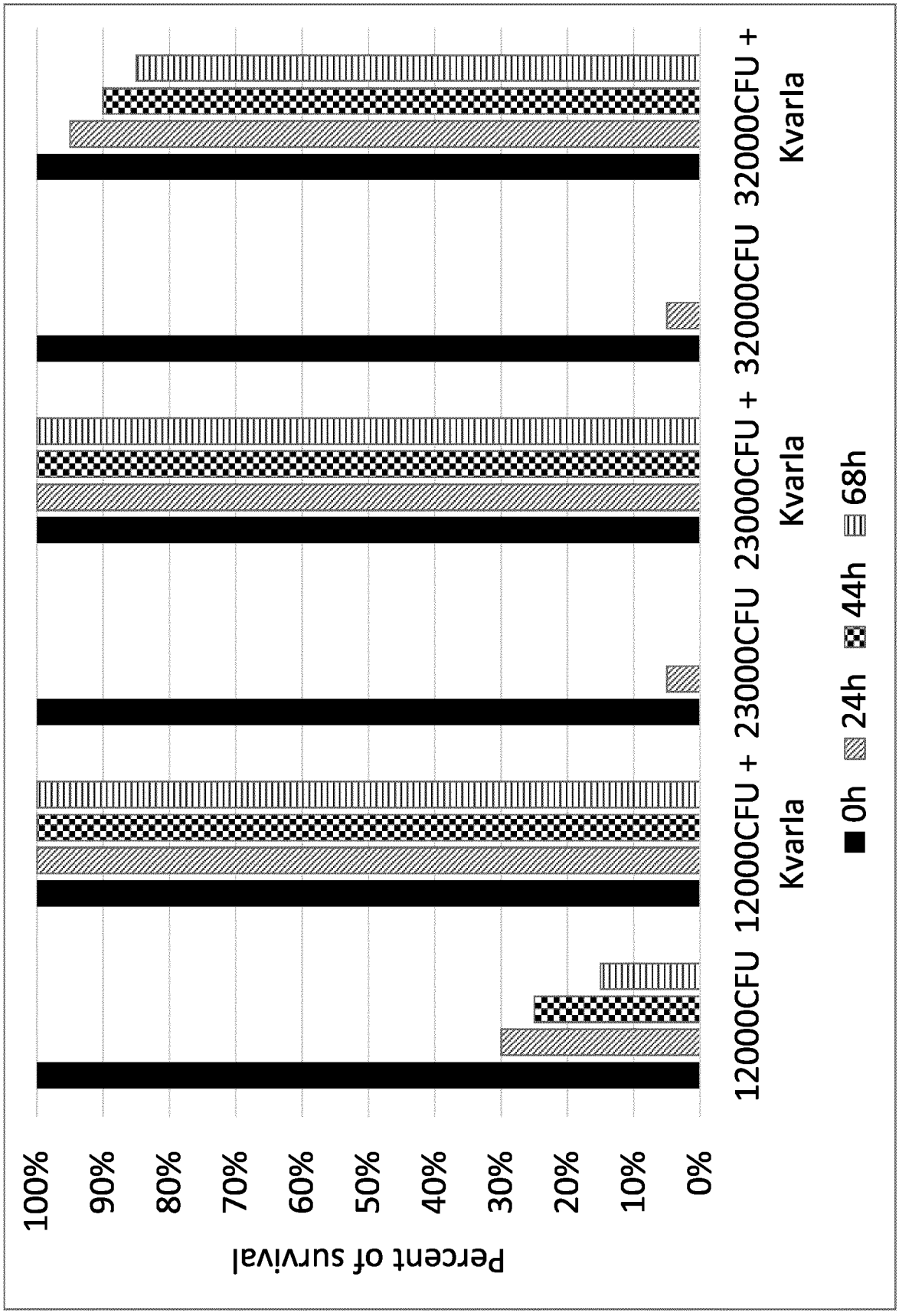

Percent of survival
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
12000CFU + 23000CFU + 32000CFU Kvarla
23000CFU + 32000CFU Kvarla
32000CFU + Kvarla
0h
24h
44h
68h

Fig. 10A

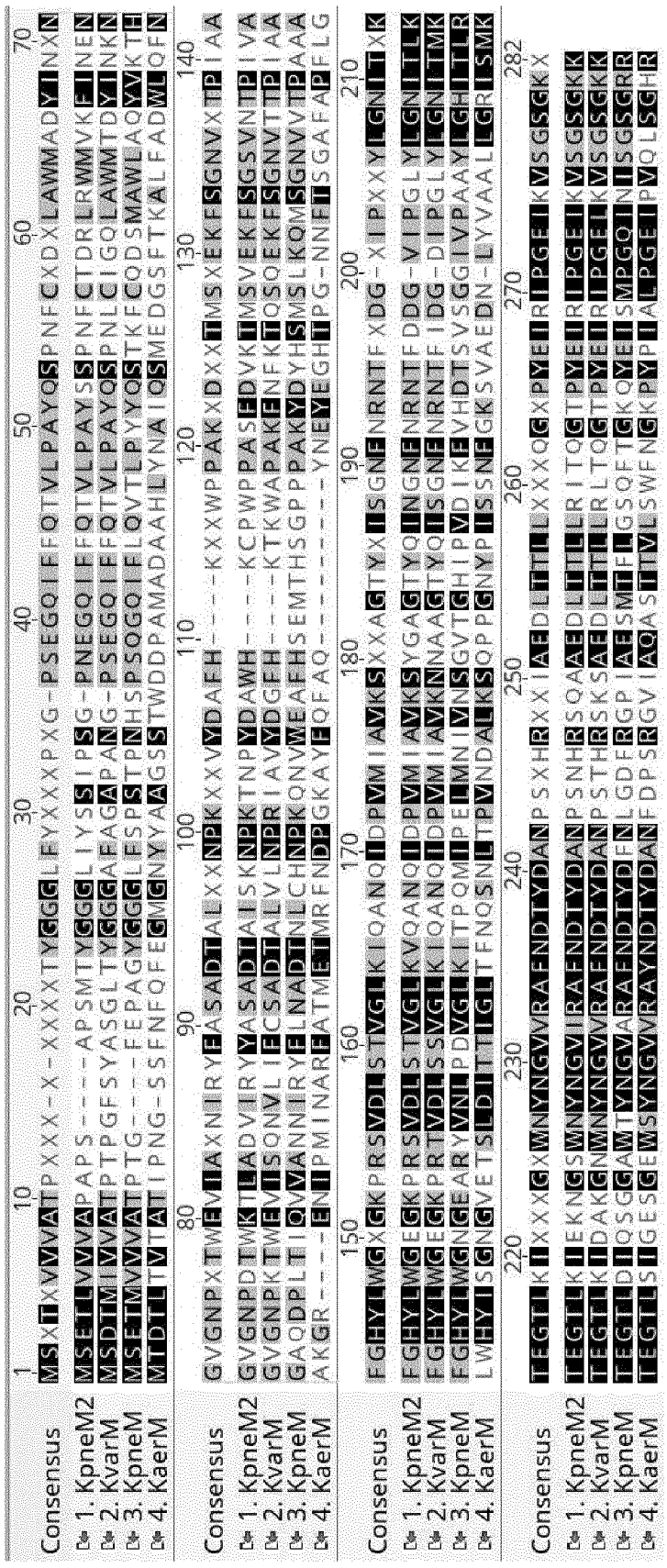
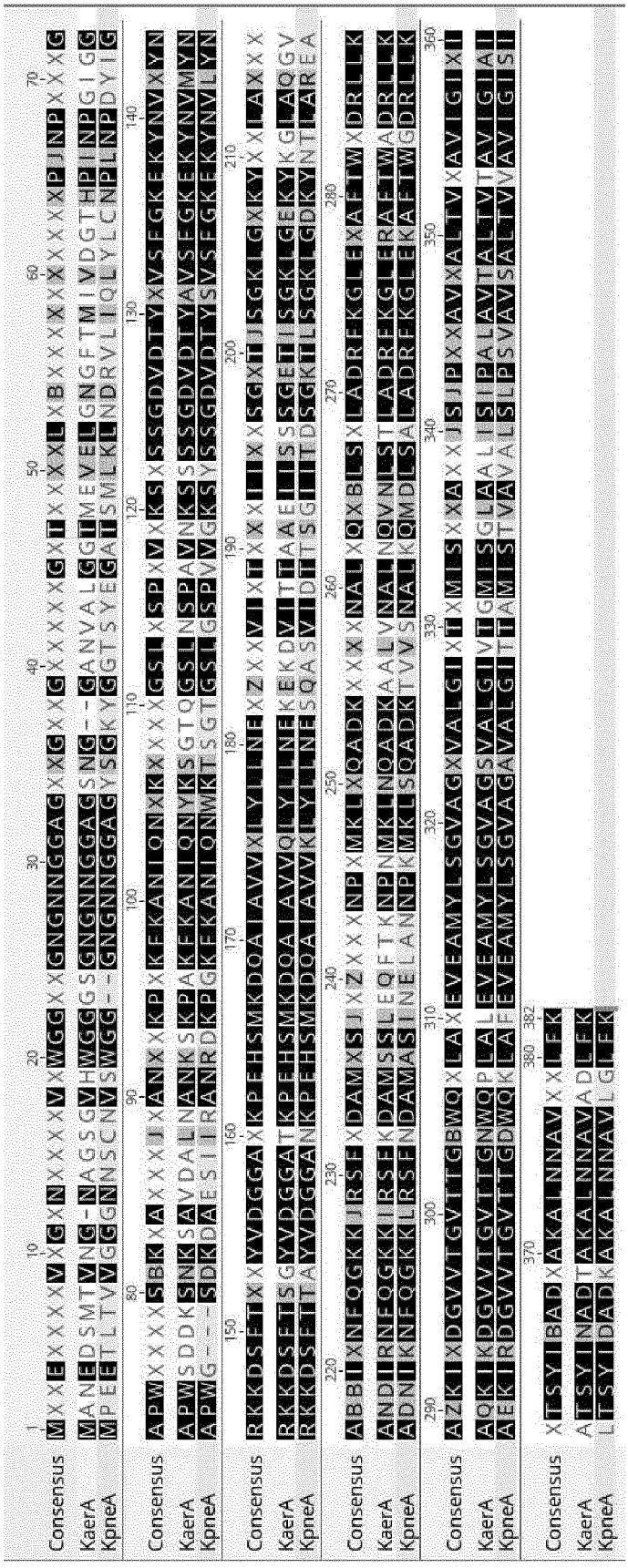


Fig. 10B



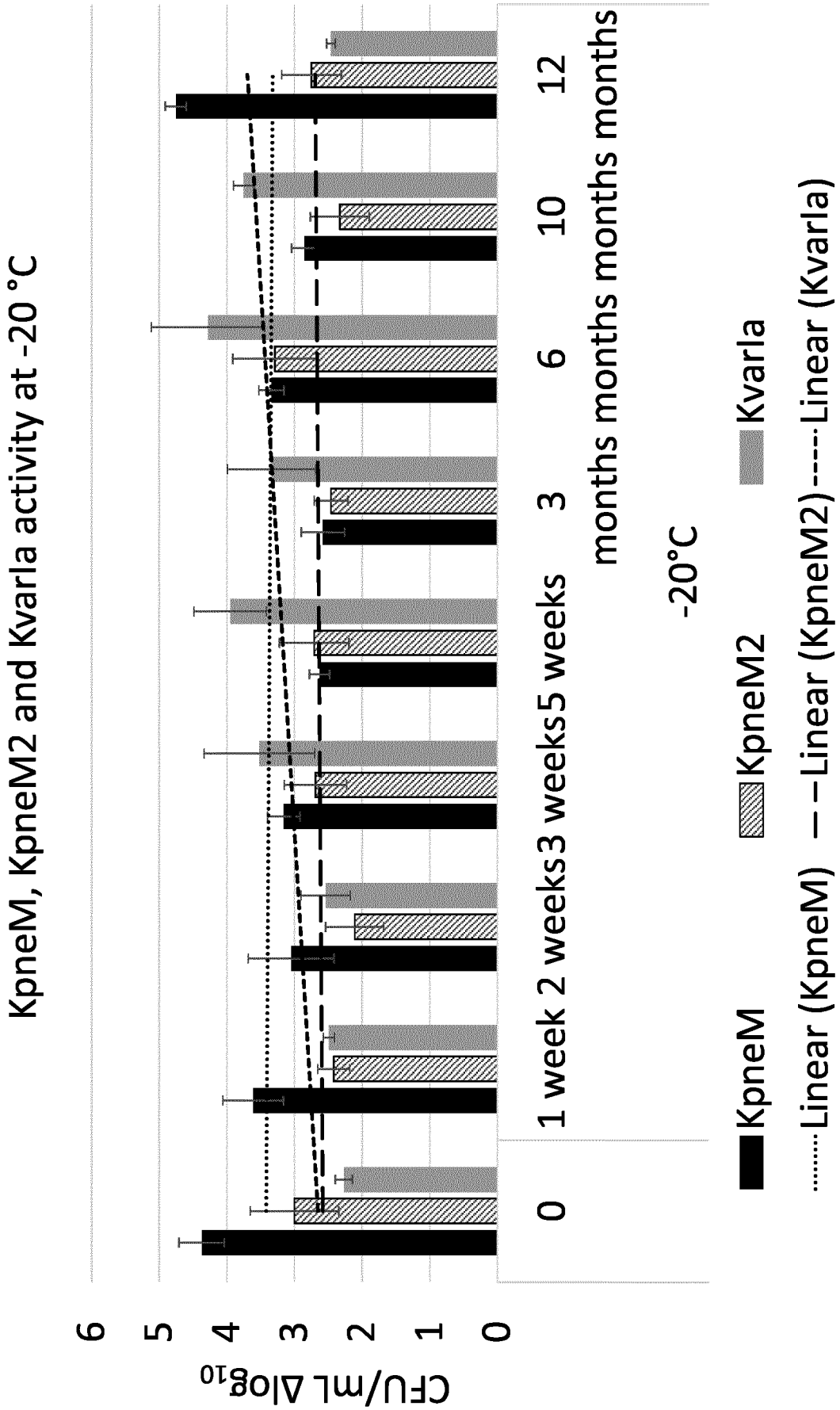

KpneM, KpneM2 and Kvarla activity at -20 °C
CFU/mL Alog10
0
1
2
3
4
5
6
0
1 week
2 weeks
3 weeks
5 weeks
3
6
10
12
months
months
months
months
-20°C
KpneM
KpneM2
Kvarla
Linear (KpneM)
Linear (KpneM2)
Linear (Kvarla)

**Fig. 11A**

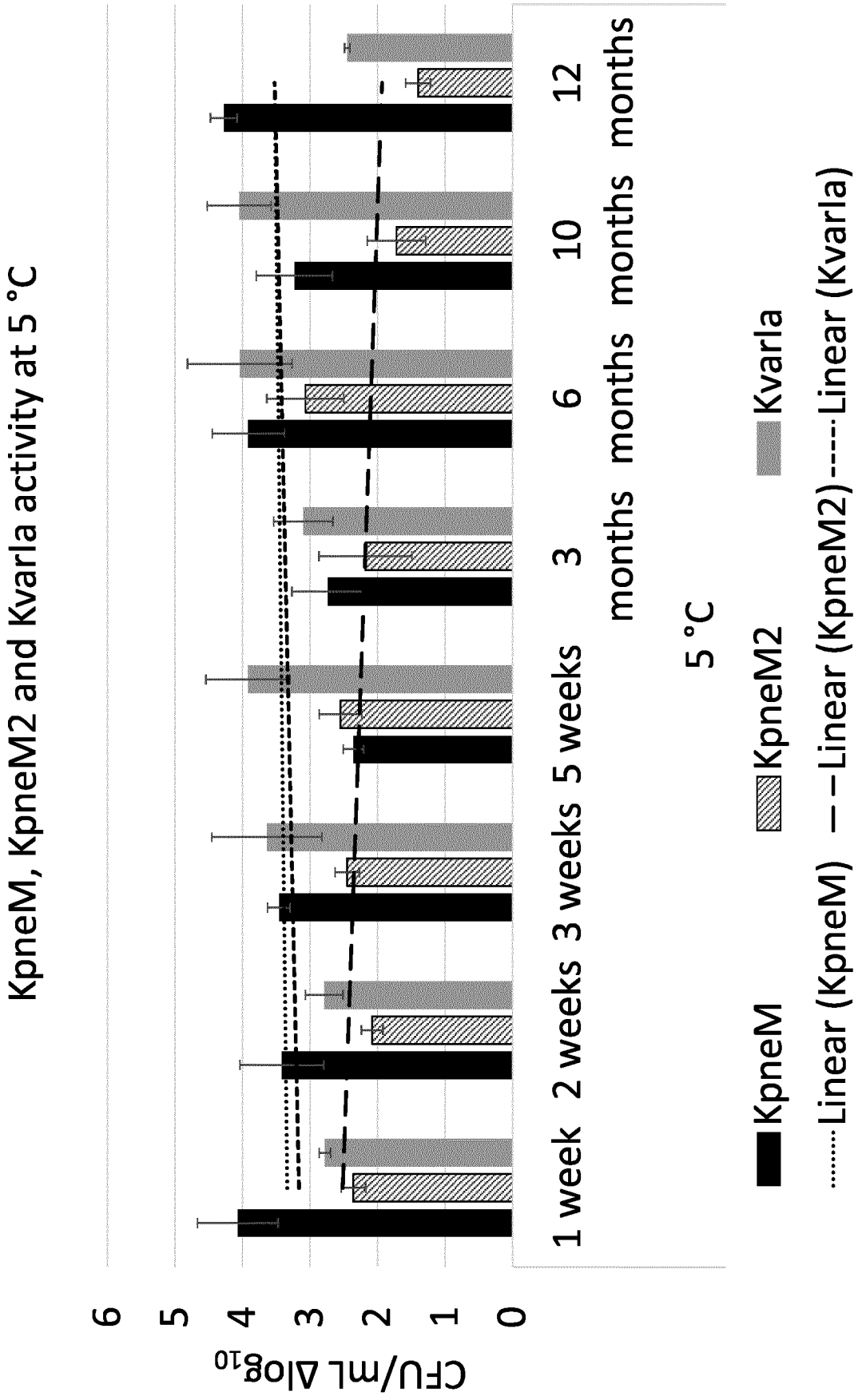

KpneM, KpneM2 and Kvarla activity at 5 °C
CFU/mL Alog10
0
1
2
3
4
5
6
1 week
2 weeks
3 weeks
5 weeks
3 months
6 months
10 months
12 months
5 °C
KpneM
KpneM2
Kvarla
Linear (KpneM)
Linear (KpneM2)
Linear (Kvarla)

**Fig. 11B**

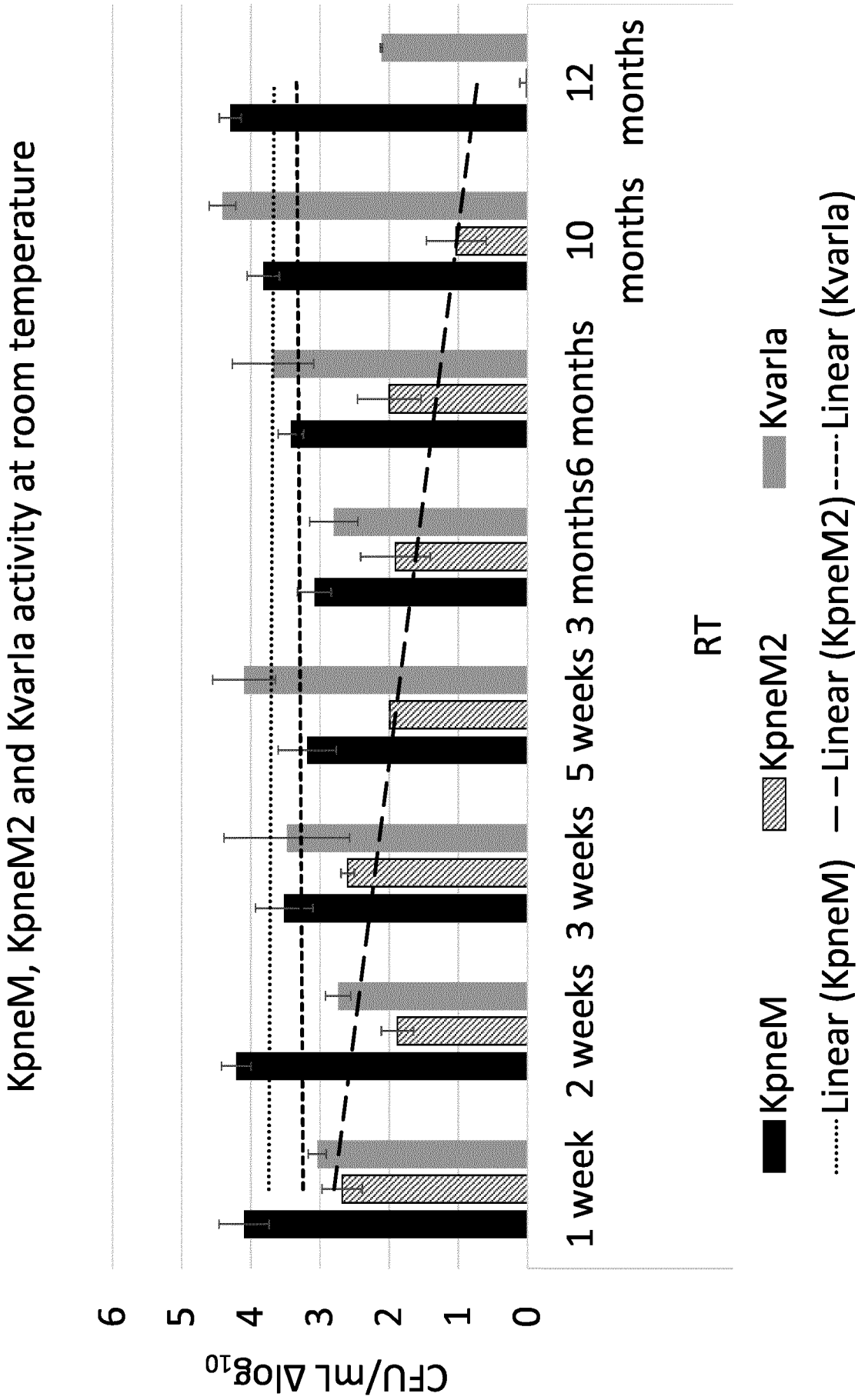

KpneM, KpneM2 and Kvarla activity at room temperature
CFU/mL Alog10
0
1
2
3
4
5
6
1 week
2 weeks
3 weeks
5 weeks
3 months
6 months
10 months
12 months
RT
KpneM
KpneM2
Kvarla
Linear (KpneM)
Linear (KpneM2)
Linear (Kvarla)

**Fig. 11C**

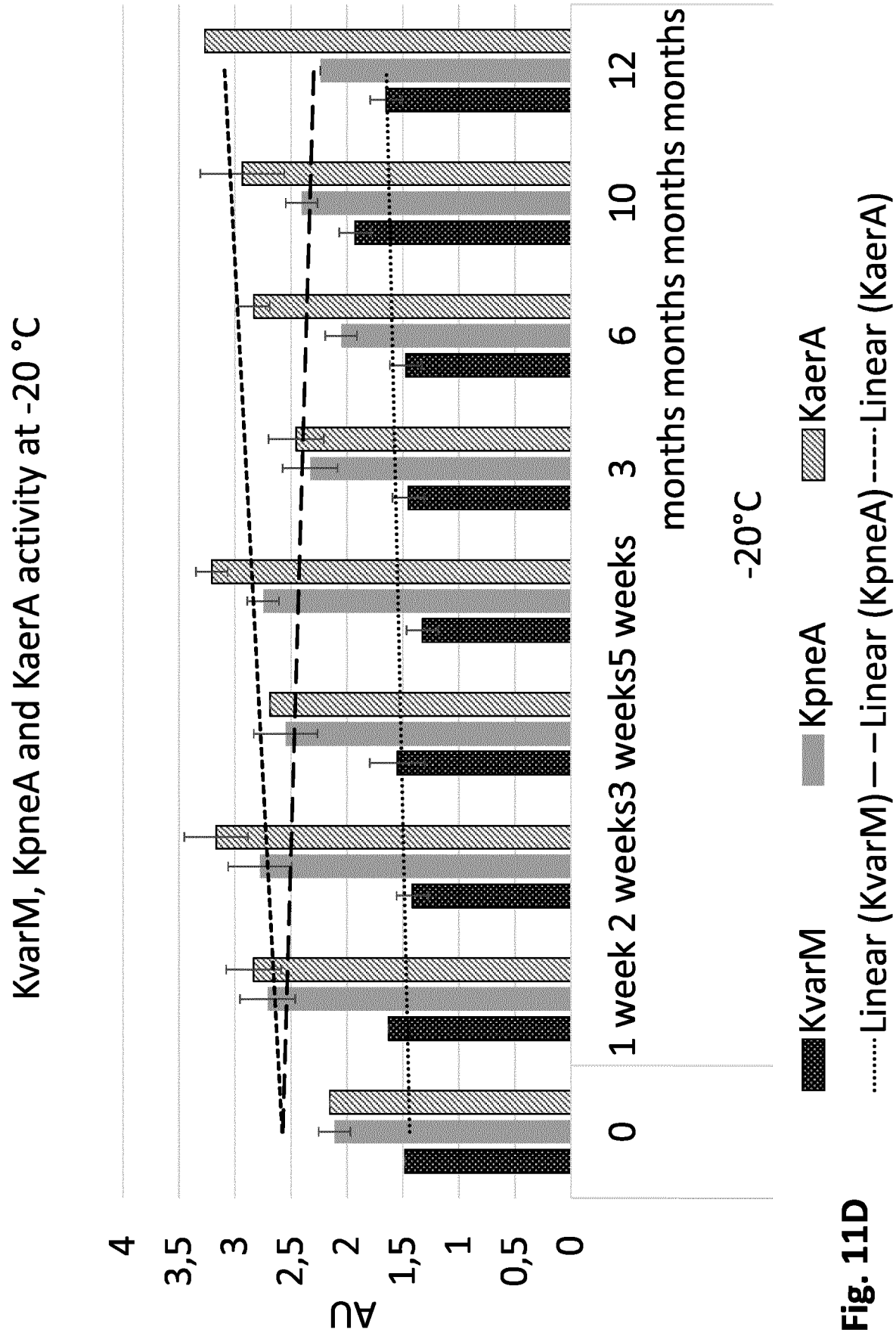

KvarM, KpneA and KaerA activity at -20 °C
AU
4
3,5
3
2,5
2
1,5
1
0,5
0
0
1 week
2 weeks
3 weeks
5 weeks
3
6
10
12
months
months
months
months
-20°C
KvarM
KpneA
KaerA
Linear (KvarM)
Linear (KpneA)
Linear (KaerA)

**Fig. 11D**

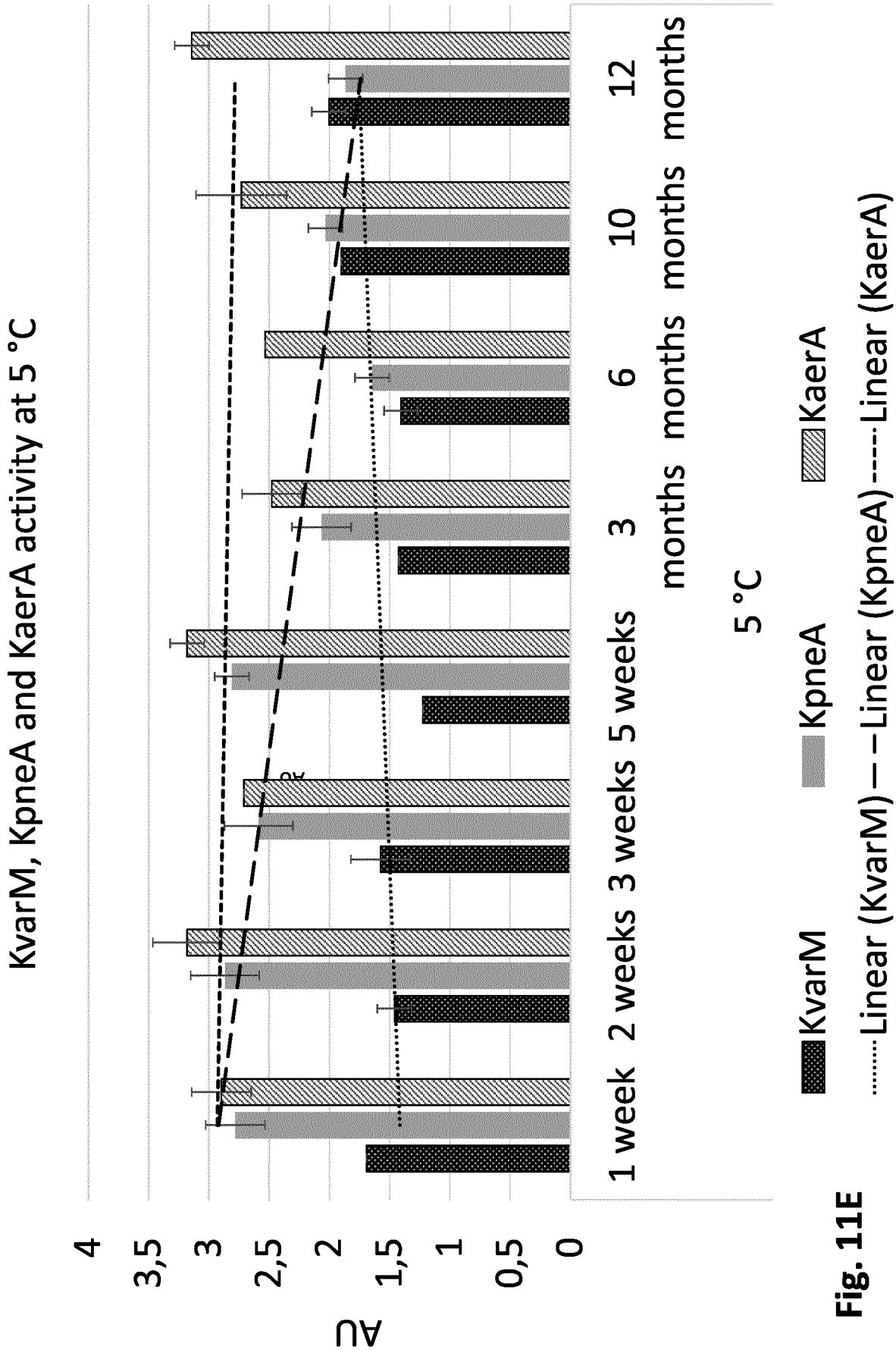

KvarM, KpneA and KaerA activity at 5 °C
AU
4
3,5
3
2,5
2
1,5
1
0,5
0
1 week
2 weeks
3 weeks
5 weeks
3 months
6 months
10 months
12 months
5 °C
KvarM
KpneA
KaerA
Linear (KvarM)
Linear (KpneA)
Linear (KaerA)

**Fig. 11E**

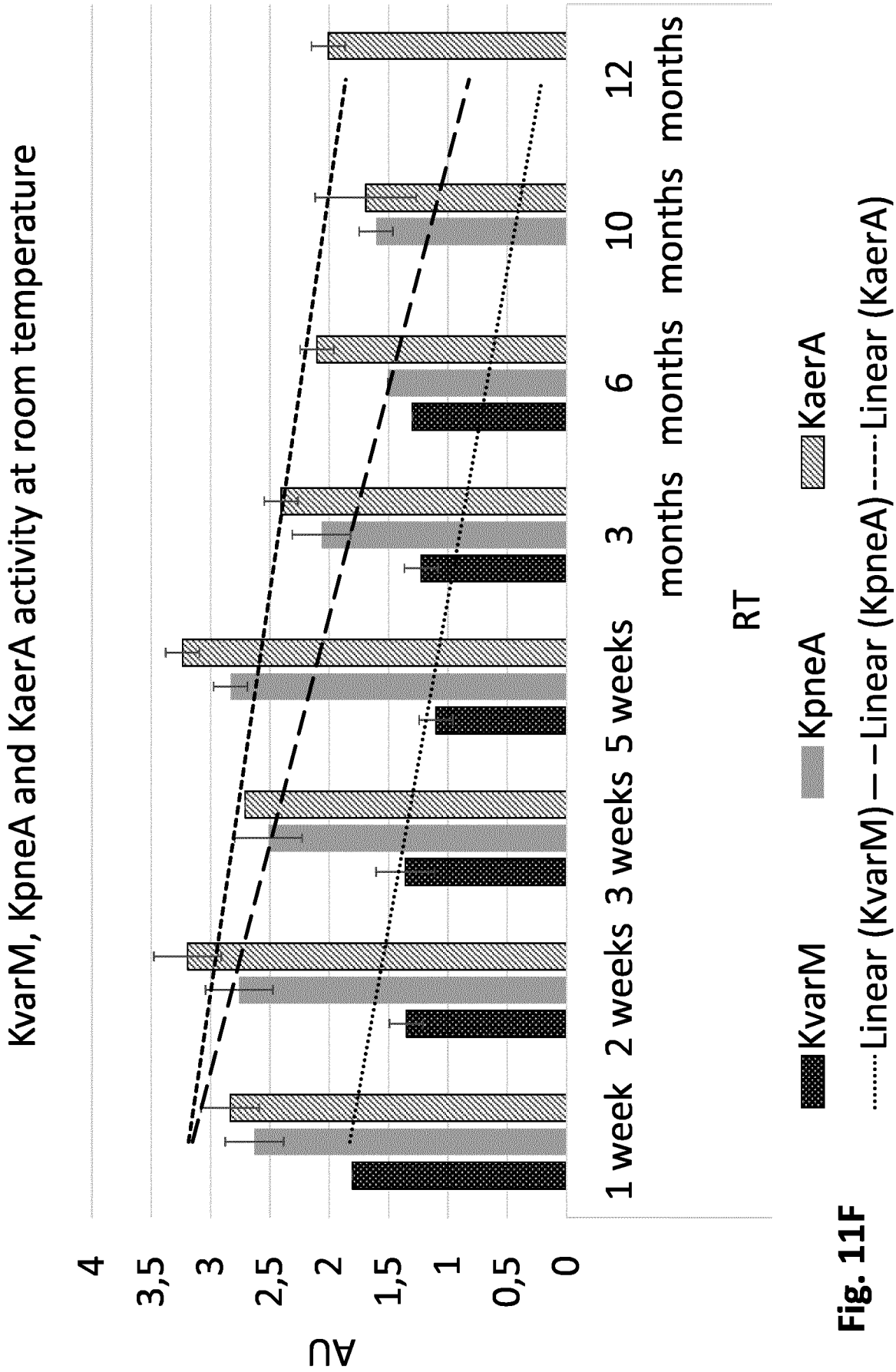

KvarM, KpneA and KaerA activity at room temperature
ΔU
4
3,5
3
2,5
2
1,5
1
0,5
0
1 week
2 weeks
3 weeks
5 weeks
3 months
6 months
10 months
12 months
RT
KvarM
KpneA
KaerA
Linear (KvarM)
Linear (KpneA)
Linear (KaerA)

**Fig. 11F**

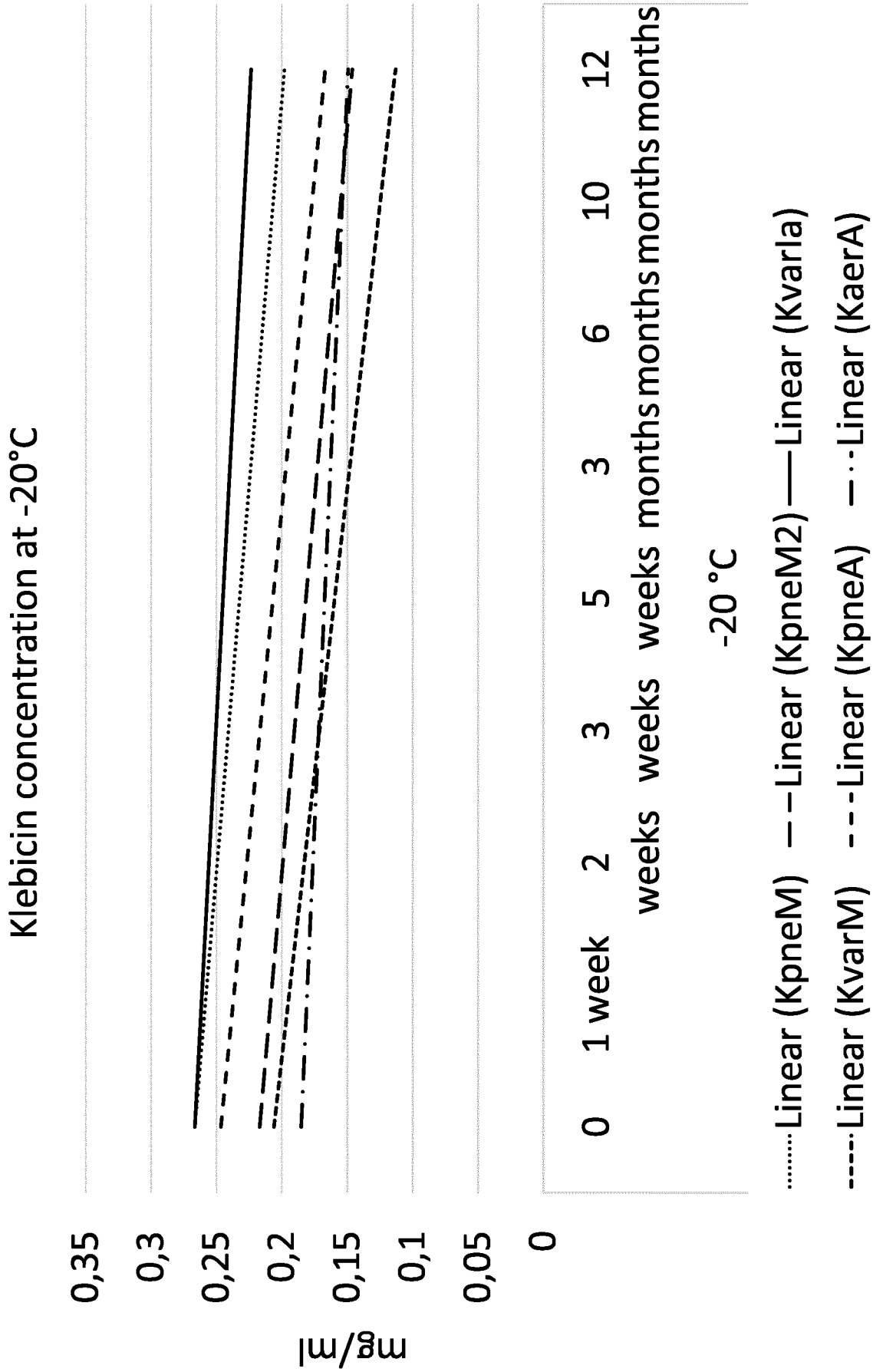

Klebicin concentration at -20°C
mg/ml
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
0 1 week 2 3 5 3 6 10 12
weeks weeks weeks months months months months
-20 °C
Linear (KpneM)
Linear (KpneM2)
Linear (Kvarla)
Linear (KvarM)
Linear (KpneA)
Linear (KaerA)

**Fig. 11G**

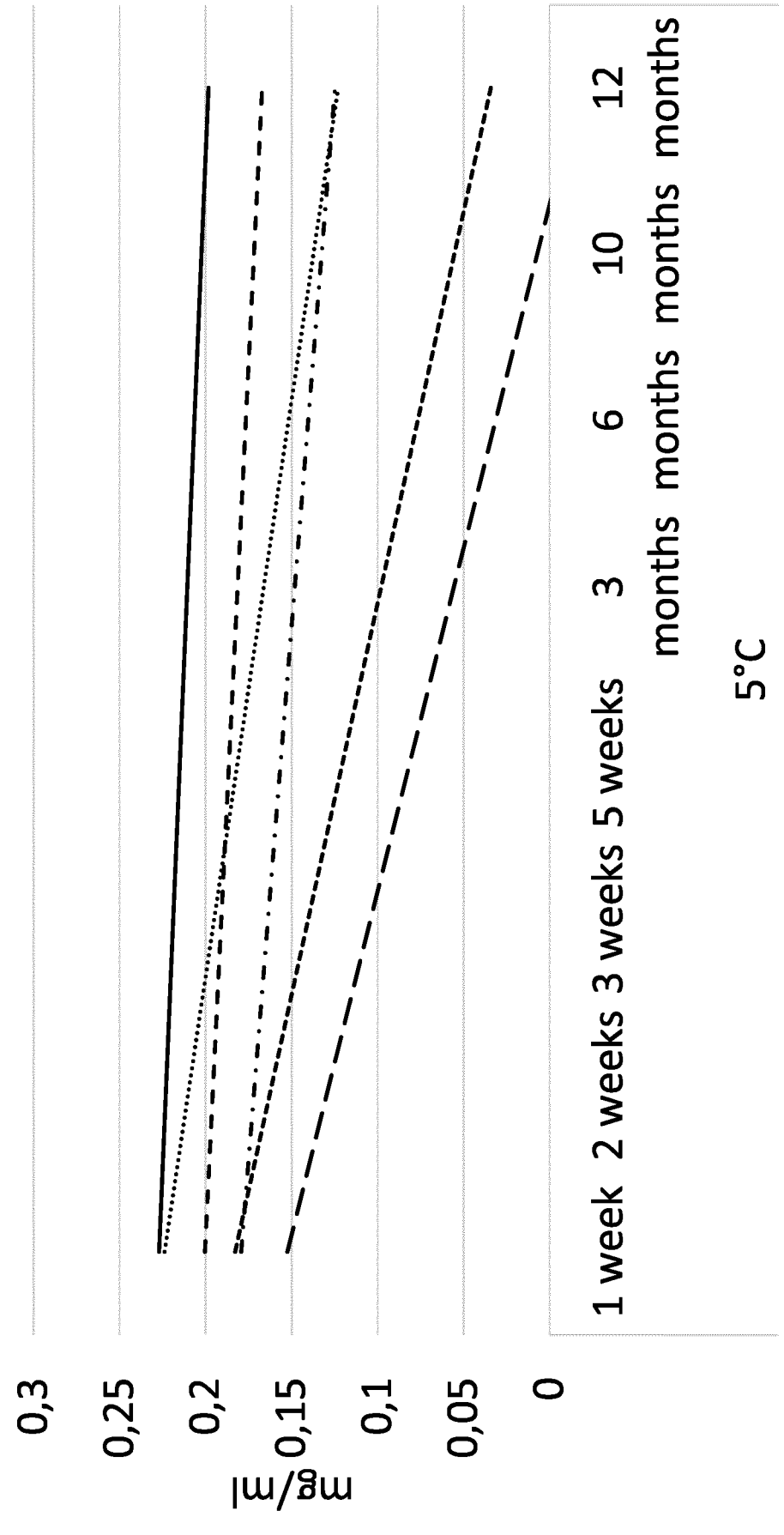

Klebicin concentration at 5 °C
mg/ml
0,3
0,25
0,2
0,15
0,1
0,05
0
1 week 2 weeks 3 weeks 5 weeks 3 6 10 12
months months months months months
5°C
Linear (KpneM)
Linear (KpneM2)
Linear (Kvarla)
Linear (KvarM)
Linear (KpneA)
Linear (KaerA)

**Fig. 11H**

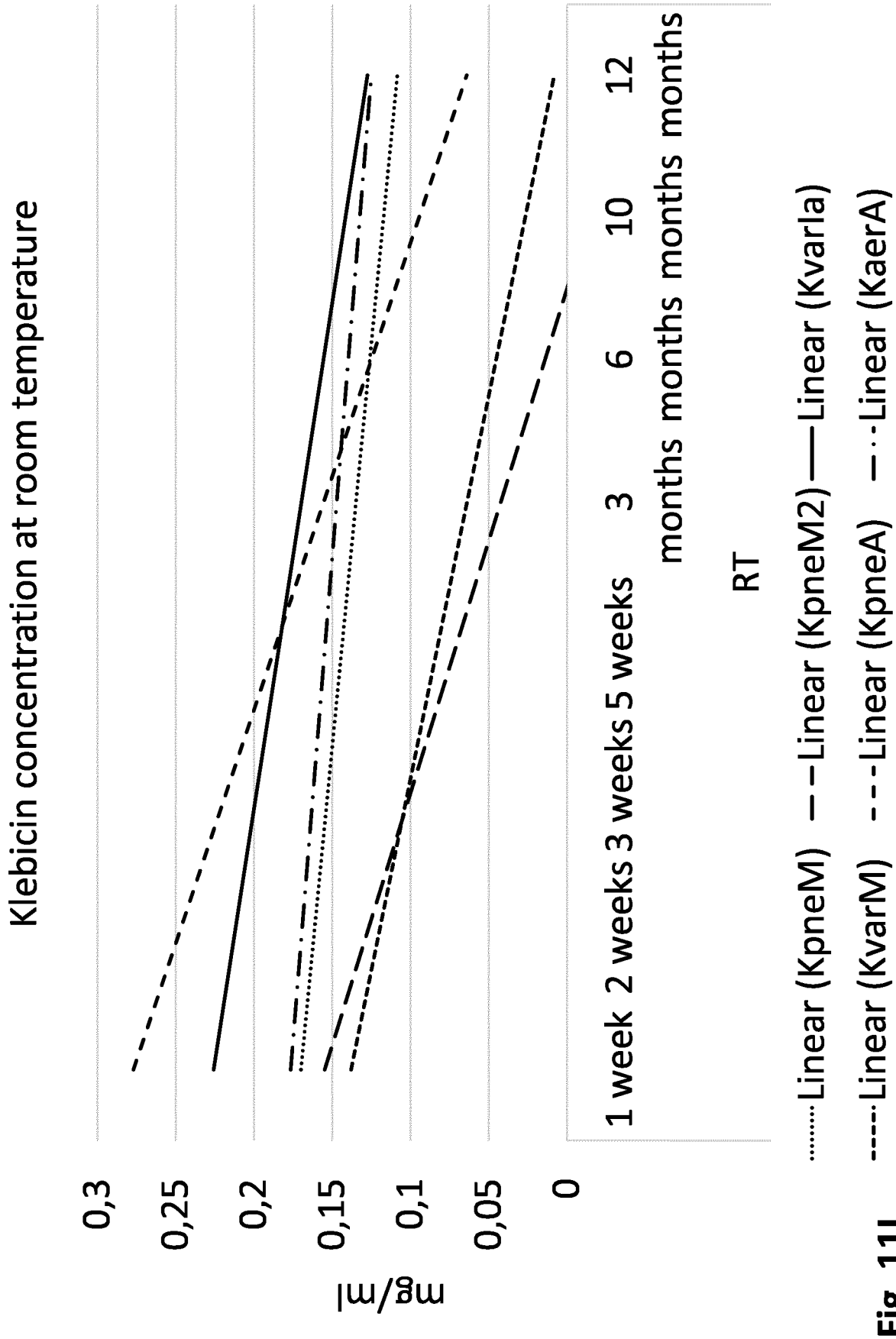

Klebicin concentration at room temperature
mg/ml
0,3
0,25
0,2
0,15
0,1
0,05
0
1 week 2 weeks 3 weeks 5 weeks 3 6 10 12
months months months months
RT
Linear (KpneM)
Linear (KpneM2)
Linear (Kvarla)
Linear (KvarM)
Linear (KpneA)
Linear (KaerA)

**Fig. 11I**

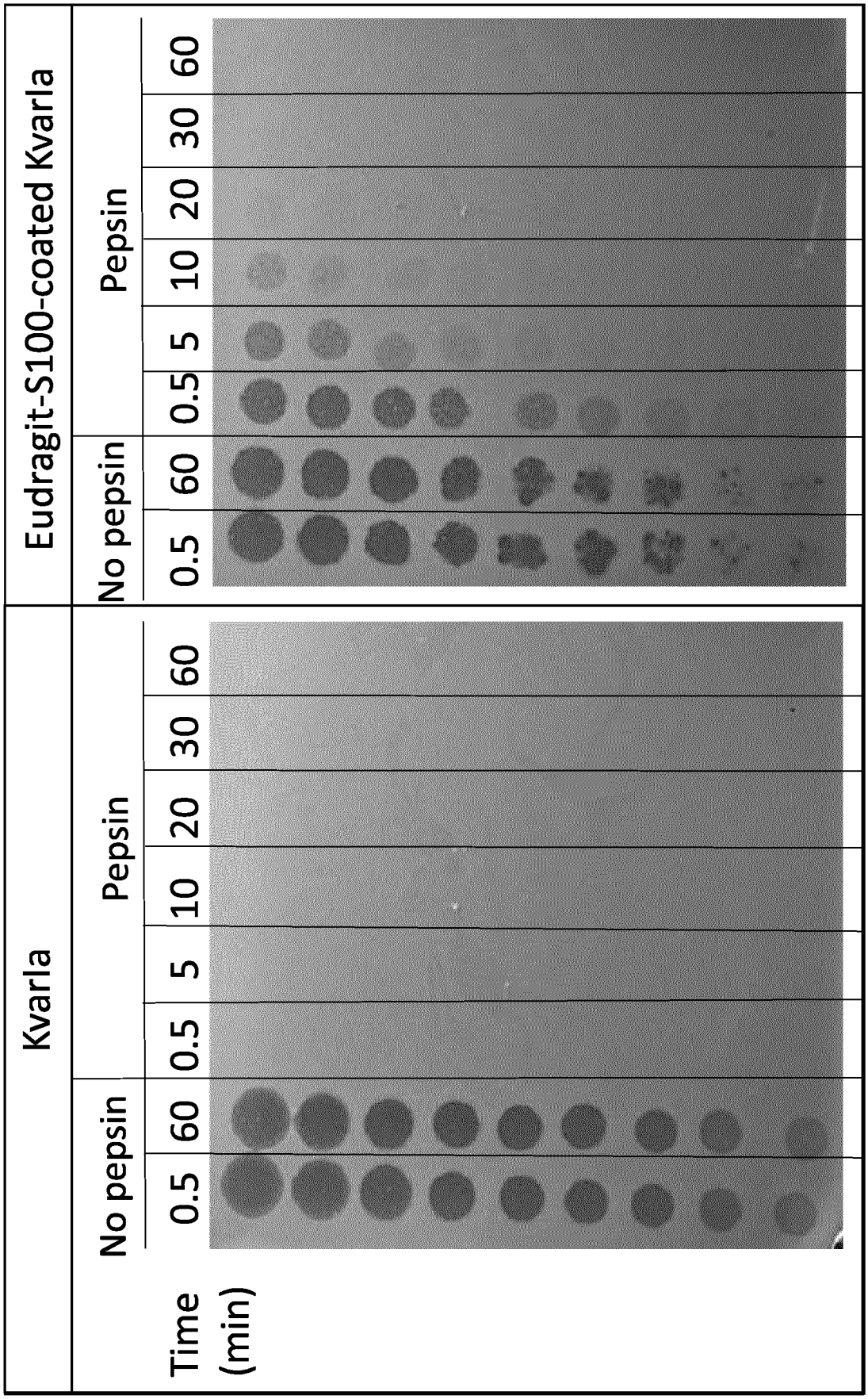

Kvarla
Eudragit-S100-coated Kvarla
No pepsin
Pepsin
Time (min)
0.5
60
0.5
5
10
20
30
60

Fig. 12A

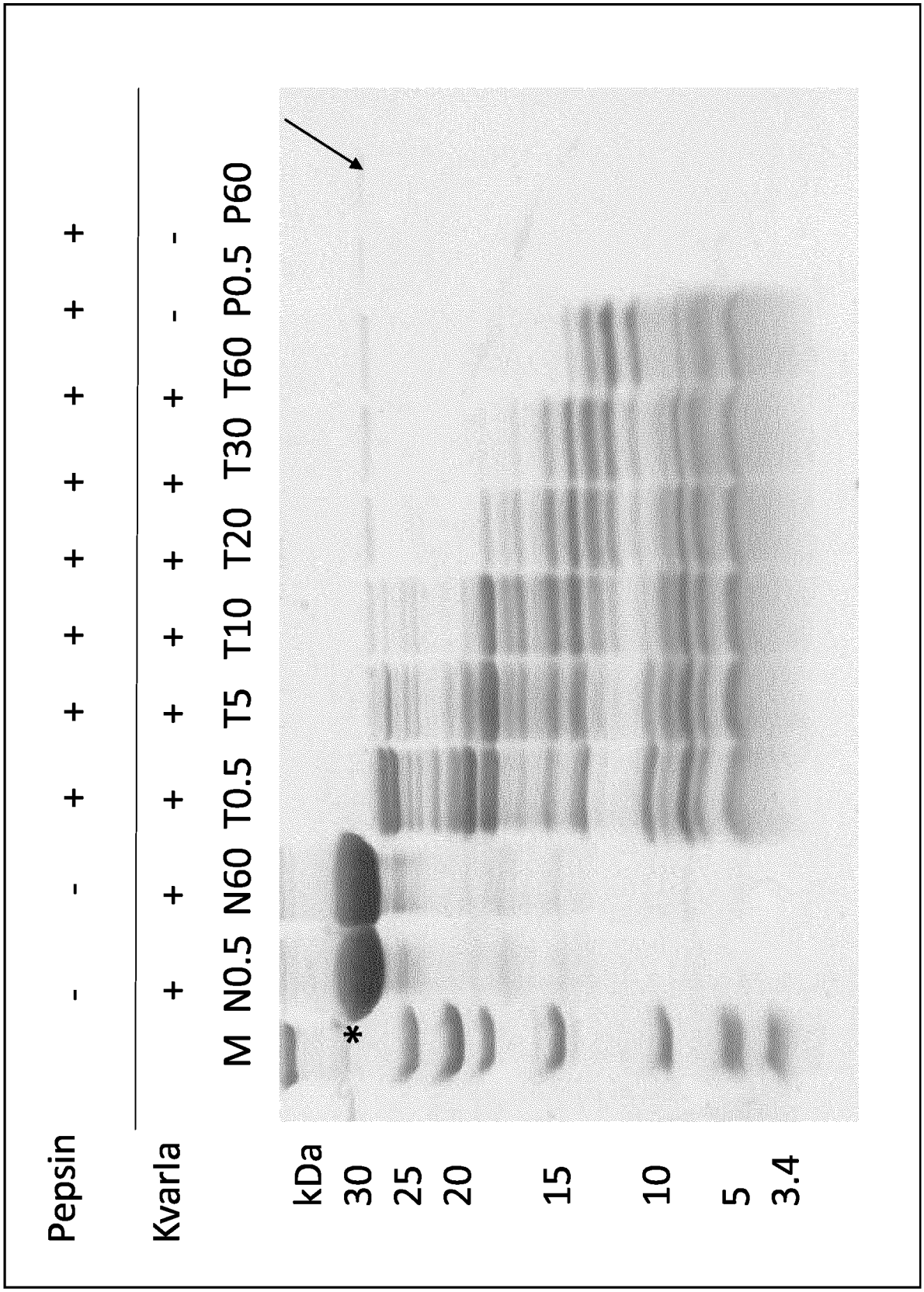

Pepsin - - + + + + + + + +
Kvarla + + + + + + + + - -
M N0.5 N60 T0.5 T5 T10 T20 T30 T60 P0.5 P60
kDa
30
25
20
15
10
5
3.4
*

Fig. 12B

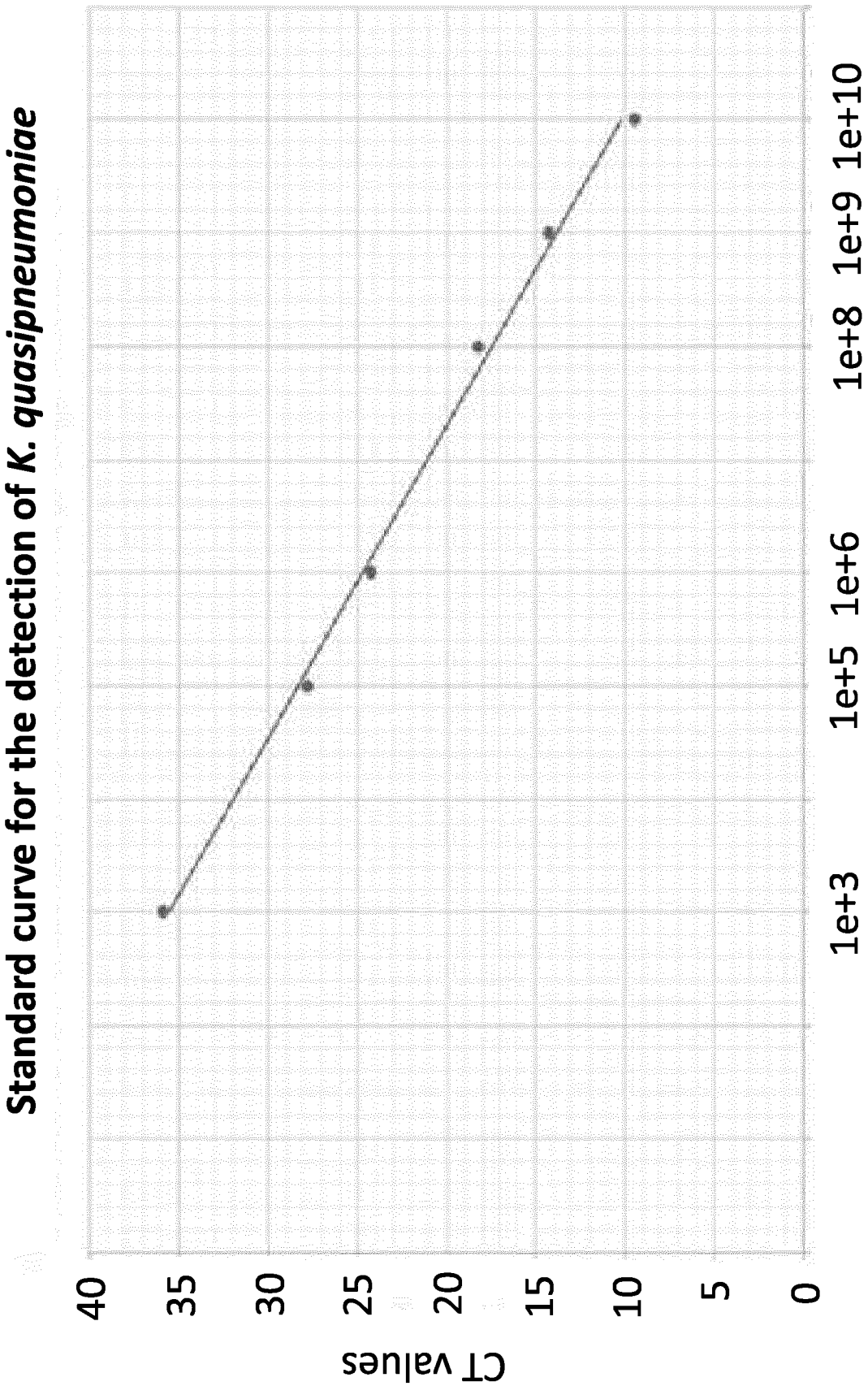

Standard curve for the detection of K. quasipneumoniae
CT values
40
35
30
25
20
15
10
5
0
1e+3
1e+5
1e+6
1e+8
1e+9
1e+10
Number of bacteria

Fig. 13

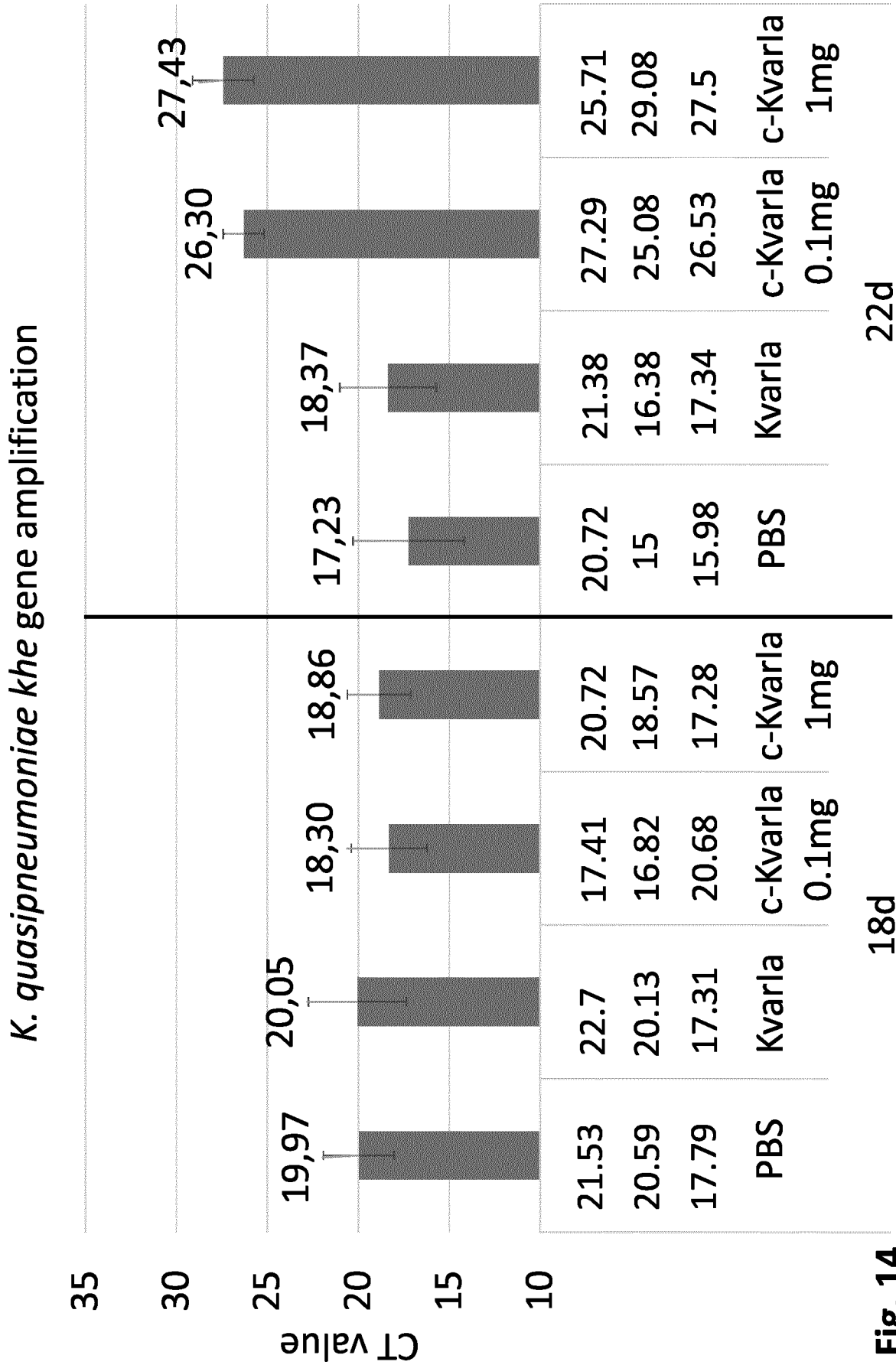

K. quasipneumoniae khe gene amplification
CT value
35
30
25
20
15
10
19,97
20,05
18,30
18,86
17,23
18,37
26,30
27,43
21.53
20.59
17.79
PBS
22.7
20.13
17.31
Kvarla
17.41
16.82
20.68
c-Kvarla 0.1mg
20.72
18.57
17.28
c-Kvarla 1mg
18d
20.72
15
15.98
PBS
21.38
16.38
17.34
Kvarla
27.29
25.08
26.53
c-Kvarla 0.1mg
25.71
29.08
27.5
c-Kvarla 1mg
22d

**Fig. 14**

# INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2020/065652

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/435
ADD. C07K14/43

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | PRASANT KUMAR JENA ET AL: "Bacteriocin PJ4 Active Against Enteric Pathogen Produced by Lactobacillus helveticus PJ4 Isolated from Gut Microflora of Wistar Rat (Rattus norvegicus): Partial Purification and Characterization of Bacteriocin", APPLIED BIOCHEMISTRY AND BIOTECHNOLOGY, vol. 169, no. 7, 1 February 2013 (2013-02-01), pages 2088-2100, XP055723662, New York ISSN: 0273-2289, DOI: 10.1007/s12010-012-0044-7 | 1-7,15, 17-21, 23,25,27 |
| A | page 2095, paragraph 2; figure 2; table 1 | 8-14,16, 22,24, 26,28 |

-----

-/--

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 20 August 2020 | 31/08/2020 |

| Name and mailing address of the ISA/ | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040,<br>Fax: (+31-70) 340-3016 | Stoyanov, Borislav |

1

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | VIDHYASAGAR VENKATASUBRAMANIAN ET AL: "Bacteriocin activity against various pathogens produced by Pediococcus pentosaceus VJ13 isolated from Idly batter", BIOMETRICAL JOURNAL, , vol. 27, no. 11 11 June 2013 (2013-06-11), pages 1497-1502, XP009522331, ISSN: 1099-0801, DOI: 10.1002/BMC.2948 Retrieved from the Internet: URL:https://onlinelibrary.wiley.com/doi/epdf/10.1002/bmc.2948 [retrieved on 2013-06-11] | 1-7,15, 17-21, 23,25,27 |
| A | figure 2; table 2 | 8-14,16, 22,24, 26,28 |
| | ----- | |
| X | KAMEL BENDJEDDOU ET AL: "Characterization and purification of a bacteriocin from Lactobacillus paracasei subsp. paracasei BMK2005, an intestinal isolate active against multidrug-resistant pathogens", WORLD JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY., vol. 28, no. 4, 3 December 2011 (2011-12-03), pages 1543-1552, XP055626079, GB ISSN: 0959-3993, DOI: 10.1007/s11274-011-0958-1 | 1-7,15, 17-21, 23,25,27 |
| A | figure 1 | 8-14,16, 22,24, 26,28 |
| | ----- | |
| X | ZACHARY S. ELLIOTT ET AL: "ABSTRACT", JOURNAL OF CLINICAL MICROBIOLOGY, vol. 57, no. 10, 24 July 2019 (2019-07-24), XP055663065, US ISSN: 0095-1137, DOI: 10.1128/JCM.00634-19 | 1-27 |
| A | table 1 | 28 |
| | ----- | |
| X | DATABASE UniProt [Online] 8 May 2019 (2019-05-08), "SubName: Full=Lipid II-degrading bacteriocin {ECO:0000313¦EMBL:EAB2325411.1};", XP002792847, retrieved from EBI accession no. UNIPROT:A0A3W8ZTQ0 Database accession no. A0A3W8ZTQ0 | 26,27 |
| A | the whole document | 1-25,28 |
| | ----- | |
| | -/-- | |

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | DATABASE UniProt [Online]<br><br>25 October 2017 (2017-10-25),<br>"RecName: Full=CHANNEL_COLICIN domain-containing protein {ECO:0000259¦Pfam:PF01024};",<br>XP002794555,<br>retrieved from EBI accession no. UNIPROT:A0A237PTB8<br>Database accession no. A0A237PTB8 | 26,27 |
| A | abstract; sequence | 1-25,28 |
| | ----- | |
| X | DATABASE UniProt [Online]<br><br>18 July 2018 (2018-07-18),<br>"RecName: Full=CHANNEL_COLICIN domain-containing protein {ECO:0000259¦Pfam:PF01024};",<br>XP002794556,<br>retrieved from EBI accession no. UNIPROT:A0A2U4DYM9<br>Database accession no. A0A2U4DYM9 | 26,27 |
| A | abstract; sequence | 1-25,28 |
| | ----- | |
| X | DATABASE UniProt [Online]<br>Iwase T., Ogura Y., Ishiwata K., Hayashi T., Yoneda M., Mizunoe Y.;: ""Complete genome sequence of Klebsiella pneumoniae YH43."",<br>XP002794557,<br>Database accession no. A0A0M4UN41 | 26,27 |
| A | abstract | 1-25,28 |
| | ----- | |
| T | ERNA DENKOVSKIENE ET AL: "Broad and Efficient Control of Klebsiella Pathogens by Peptidoglycan-Degrading and Pore-Forming Bacteriocins Klebicins",<br>SCIENTIFIC REPORTS,<br>vol. 9, no. 1,<br>28 October 2019 (2019-10-28), XP055723685,<br>DOI: 10.1038/s41598-019-51969-1<br>the whole document | |
| | ----- | |

1